

GenCore version 5.1.3  
Copyright (c) 1993 - 2003 CompuGen Ltd.

## OM nucleic - nucleic search, using sw model

Run on: February 3, 2003, 14:22:43 ; Search time 3867 Seconds  
(without alignments)  
10799.722 Million cell updates/sec

Title: US-09-763-748-1  
Sequence: 1435  
1 ctggcgcgcgcgtgggaacca.....gataataaglatracacgg 1435

Scoring table: IDENTITY\_NUC  
Gapop 10.0 , Gapept 1.0

Searched: 2054640 seqs, 14551402878 residues

Total number of hits satisfying chosen parameters: 4109280

Minimum DB seq length: 0  
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%  
Maximum Match 100%

Listing first 45 summaries

Database : GenEmbl:\*

1: gb\_ba:\*

2: gb\_hcg:\*

3: gb\_in:\*

4: gb\_om:\*

5: gb\_ov:\*

6: gb\_pat:\*

7: gb\_ph:\*

8: gb\_pl:\*

9: gb\_pr:\*

10: gb\_ro:\*

11: gb\_sts:\*

12: gb\_sy:\*

13: gb\_un:\*

14: gb\_vi:\*

15: em\_ba:\*

16: em\_fun:\*

17: em\_hum:\*

18: em\_in:\*

19: em\_mu:\*

20: em\_om:\*

21: em\_or:\*

22: em\_ov:\*

23: em\_pat:\*

24: em\_ph:\*

25: em\_pl:\*

26: em\_ro:\*

27: em\_sts:\*

28: em\_un:\*

29: em\_vi:\*

30: em\_hcg\_hum:\*

31: em\_hcg\_inv:\*

32: em\_hcg\_other:\*

33: em\_hcg\_mus:\*

34: em\_hcg\_pln:\*

35: em\_hcg\_rtd:\*

36: em\_hcg\_man:\*

37: em\_hcg\_vrt:\*

38: em\_sy:\*

39: em\_hcgo\_hum:\*

40: em\_hcgo\_mus:\*

41: em\_hcgo\_other:\*

score greater than or equal to the score of the result being printed,  
and is derived by analysts of the total score distribution.

## SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	1435	100.0	1435	6 AR098746	AR098746 Sequence
2	1435	100.0	1435	6 HUMTRADD	LA1690 Homo sapien
3	1435	100.0	1441	6 I27061	I27061 Sequence 1
4	1433.4	99.9	1475	9 BC004491	BC004491 Homo sapi
5	1243.8	86.7	1833	9 AK090673	AK090673 Homo sapi
6	1005.8	70.1	152953	9 AC074143	AC074143 Homo sapi
7	1005.8	70.1	209574	9 AC040160	AC040160 Homo sapi
8	990.8	69.0	228516	2 AC008594	AC008594 Homo sapi
9	512.2	35.7	579	2 HSAJ11616	HSAJ11616 Homo sapi
10	434.4	30.3	208936	2 AC010821	AC010821 Homo sapi
11	427.6	29.8	1384	6 I27062	I27062 Sequence 3
12	398.4	27.8	927	9 HSAJ11615	HSAJ11615 Homo sapi
13	282.6	19.7	237933	2 AC124713	AC124713 Mus muscu
14	272	19.0	180835	2 AC120484	AC120484 Rattus no
15	200	13.9	200	6 AR211525	AR211525 Sequence
16	175.8	12.3	334	10 AF517017	AF517017 Rattus no
17	175.4	12.2	740	9 HSAJ31001	AJ331001 Homo sapi
18	175.4	12.2	743	9 HSAJ36488	AJ334498 Homo sapi
19	172.2	12.0	1184	9 HSAJ34498	AJ334498 Homo sapi
20	165.8	11.6	731	9 HSAJ24630	AJ324630 Homo sapi
21	164.6	11.5	704	9 HSAJ11614	AJ311614 Homo sapi
22	127	8.9	552	10 BC028542	BC028542 Mus muscu
23	111.4	7.8	882	5 AF231014	AF231014 Danio rer
24	103.6	7.2	180835	2 AC120484	AC120484 Rattus no
25	77.4	5.4	155662	2 AC112388	AC112388 Rattus no
26	72.8	5.1	3060	3 AF027735	AF027735 Nephila c
27	72	5.0	125020	9 AF429315	AF429315 Homo sapi
28	70.6	4.9	125020	9 AF429315	AF429315 Homo sapi
29	69.8	4.9	1279	11 PM2H12G	AL684840 Penicillii
30	69.4	4.8	40476	1 SCCT75A	AL133220 Streptomy
31	69.2	4.8	2793	6 AR009990	AR009990 Sequence
32	69.2	4.8	2793	6 I95876	I95876 Sequence 1
33	65.2	4.5	34023	1 SC262	AL445963 Streptomy
34	63.8	4.4	1150	14 HSAJULR3	J02079 Epstein-Bar
35	63.8	4.4	1926	6 AX107940	AX107940 Sequence
36	63.8	4.4	2580	6 AR108994	AR108994 Sequence
37	63.8	4.4	5452	6 AR083151	AR083151 Sequence
38	63.8	4.4	5452	12 U02454	U02454 Cloning vec
39	63.8	4.4	9600	6 A92665	A92665 Sequence 1
40	63.8	4.4	10596	6 AR158345	AR158345 Sequence
41	63.8	4.4	10596	6 I25041	I25041 Sequence 15
42	63.8	4.4	10596	6 I30503	I30503 Sequence 15
43	63.8	4.4	10737	12 XXU02428	U02428 Cloning vec
44	63.8	4.4	10850	12 U02455	U02455 Cloning vec
45	63.8	4.4	172281	14 EBV	V01555 Epstein-Bar

## ALIGNMENTS

RESULT 1

AR098746 1435 bp DNA linear PAT 14-FEB-2001

LOCUS AR098746

DEFINITION Sequence 1 from patent US 6077672.

ACCESSION AR098746

VERSION AR098746.1 GI:12808512

KEYWORDS

SOURCE Unknown.

ORGANISM Unknown.

REFERENCE 1 (bases 1 to 1435)

AUTHORS Montu, B.P. and Cowser, L.M.

TITLE Antisense modulation of TRAD expression

JOURNAL Patent: US 6077672-A 1 20-JUN-2000;

FEATURES Location/Qualifiers

Pred. No. is the number of results predicted by chance to have a

source 1..1435  
/organism="unknown"  
BASE COUNT 245 a 437 c 494 g 259 t  
ORIGIN

Query Match 100.0%; Score 1435; DB 6; Length 1435;  
Best Local Similarity 100.0%; Pred. No. 2.3e-220;  
Matches 1435; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 CCGGCGGGGCTGGGAACCCAGGCCCCCGGAGGCGCCAGAGAGTGAATGGACTGGG 60  
DB 1 CCGGCGGGGCTGGGAACCCAGGCCCCCGGAGGCGCCAGAGAGTGAATGGACTGGG 60  
QY 61 CAAAATGGCAGGAAGTGGTGGCAGCGCATACCTGTTTGGAGTCTCGCTGAGC 120  
DB 61 CAAAATGGCAGGAAGTGGTGGCAGCGCATACCTGTTTGGAGTCTCGCTGAGC 120  
QY 121 AAGGTGCTCTGTGGATGCTTACGCGCACCCCGACAGAAAGTGGCAGTGAACAGGCT 180  
DB 121 AAGGTGCTCTGTGGATGCTTACGCGCACCCCGACAGAAAGTGGCAGTGAACAGGCT 180  
QY 181 CTGCAAGGCTCTGTGGATGCTTACGCGCACCCCGACAGAAAGTGGCAGTGAACAGGCT 240  
DB 181 CTGCAAGGCTCTGTGGATGCTTACGCGCACCCCGACAGAAAGTGGCAGTGAACAGGCT 240  
QY 241 CACCGCAGGACCCCGCAGCTGATCTGAGCTGCAATTCTGCGGGCGGACCCCTGTGGC 300  
DB 241 CACCGCAGGACCCCGCAGCTGATCTGAGCTGCAATTCTGCGGGCGGACCCCTGTGGC 300  
QY 301 GCGCTTCTCGCGCTTACCGCGAGAGGGGCGCTGCGCGCGCGCTGAGAGAGCCTGAGC 360  
DB 301 GCGCTTCTCGCGCTTACCGCGAGAGGGGCGCTGCGCGCGCGCTGAGAGAGCCTGAGC 360  
QY 361 GCGCGCTGCGCCAGAGACTCGTGCCTGCGCTGCACTGAGAGTGGCGCCGCGCGAGCGG 420  
DB 361 GCGCGCTGCGCCAGAGACTCGTGCCTGCGCTGCACTGAGAGTGGCGCCGCGCGAGCGG 420  
QY 421 CTGCAAGCTTGTGCGCAGCAGAGAGCGCTGTTGAGTTGCACTCTAGCCCGAGCGCC 480  
DB 421 CTGCAAGCTTGTGCGCAGCAGAGAGCGCTGTTGAGTTGCACTCTAGCCCGAGCGCC 480  
QY 481 GACCGGCTCTCGGAGTGAAGACTGCTGAGCTGAGAGATGCGCTGCGAAATCTGAAGTGC 540  
DB 481 GACCGGCTCTCGGAGTGAAGACTGCTGAGCTGAGAGATGCGCTGCGAAATCTGAAGTGC 540  
QY 541 GCGTGGGGGGCCCCGGGCTGCGCAGCGGGAGTGCCTTCCGCCCTTGAAGCCCCCGGTG 600  
DB 541 GCGTGGGGGGCCCCGGGCTGCGCAGCGGGAGTGCCTTCCGCCCTTGAAGCCCCCGGTG 600  
QY 601 CCTCTCTGTGCGAGGTGAAGCGCGCGCGCGCGCAGCTGCGCAGCTTTTCTGTTTC 660  
DB 601 CCTCTCTGTGCGAGGTGAAGCGCGCGCGCGCGCAGCTGCGCAGCTTTTCTGTTTC 660  
QY 661 CAGGGTCAAGCTGTAGTGAATCGGCGCTGAGCTGAAGAGCCAGACGTTTCCGCGCC 720  
DB 661 CAGGGTCAAGCTGTAGTGAATCGGCGCTGAGCTGAAGAGCCAGACGTTTCCGCGCC 720  
QY 721 TCTGTGGGTCTCAAAATGGCGCAAGTGGGGCGCTCACTGACAGAGAGCTGCGGGGCGCTG 780  
DB 721 TCTGTGGGTCTCAAAATGGCGCAAGTGGGGCGCTCACTGACAGAGAGCTGCGGGGCGCTG 780  
QY 781 CCGGACCCCGCGCTGAGCTGCGGCTGAGAGTGAAGAGCGAGGAGCGCGCGCTGAGAGC 840  
DB 781 CCGGACCCCGCGCTGAGCTGCGGCTGAGAGTGAAGAGCGAGGAGCGCGCGCTGAGAGC 840  
QY 841 GCGTTCAGAGCTGTGCGGCGCTTGTGAGAGCGCGAGAGCGCGCGCTGAGAGC 900  
DB 841 GCGTTCAGAGCTGTGCGGCGCTTGTGAGAGCGCGAGAGCGCGCGCTGAGAGC 900  
QY 901 CTGGTGGAGGCACTGAGAGAGAGAGCTCAAGAGCTGCGAGAGAGAGCTTGTGGGCTG 960  
DB 901 CTGGTGGAGGCACTGAGAGAGAGAGCTCAAGAGCTGCGAGAGAGAGCTTGTGGGCTG 960

QY 961 ACCGATCCCAATGAGCGGCTGAGCTAGACAGAGGCTGAGCCAGCTTTTGGAGAACTTGG 1020  
DB 961 ACCGATCCCAATGAGCGGCTGAGCTAGACAGAGGCTGAGCCAGCTTTTGGAGAACTTGG 1020  
QY 1021 ATGGCTTTAGGCTTCTTCTGCGGCTATTGCTGAACCCCTGTTCATTCACGGGACCTTGA 1080  
DB 1021 ATGGCTTTAGGCTTCTTCTGCGGCTATTGCTGAACCCCTGTTCATTCACGGGACCTTGA 1080  
QY 1081 AACTCCACTTGGCTTATGCTGAGCCTGCTGGGGCAGAGTTGATGCTTCCCGAGAG 1140  
DB 1081 AACTCCACTTGGCTTATGCTGAGCCTGCTGGGGCAGAGTTGATGCTTCCCGAGAG 1140  
QY 1141 CCAGACCACTGGGGGTGATCATTTGGGAGATTCTGCTCAGGACTTTGATAGAGTGGG 1200  
DB 1141 CCAGACCACTGGGGGTGATCATTTGGGAGATTCTGCTCAGGACTTTGATAGAGTGGG 1200  
QY 1201 GTGGGGGGGACTTGTCTTTGAGATCAGCTCACTTTCCTCCATCCAGAGCGGGCTTGA 1260  
DB 1201 GTGGGGGGGACTTGTCTTTGAGATCAGCTCACTTTCCTCCATCCAGAGCGGGCTTGA 1260  
QY 1261 CAGCAGCGCTTACAGTTTCACTCACTGAAGCACTTGAATCTTGTGCTTGGACTTCAAT 1320  
DB 1261 CAGCAGCGCTTACAGTTTCACTCACTGAAGCACTTGAATCTTGTGCTTGGACTTCAAT 1320  
QY 1321 CTTGGGTGCTGAGATCTGAGTGAATGAAGAGATGATCTTGTCTGCGCCAGC 1380  
DB 1321 CTTGGGTGCTGAGATCTGAGTGAATGAAGAGATGATCTTGTCTGCGCCAGC 1380  
QY 1381 TCACACTCAGCGTGGAGCCCGGATGTTTAAGCATGATTAATTAATTAACAGG 1435  
DB 1381 TCACACTCAGCGTGGAGCCCGGATGTTTAAGCATGATTAATTAATTAACAGG 1435

## RESULT 2

HUMTRADD 1435 bp mRNA linear PRI 26-APR-1996  
LOCUS Homo sapiens TNF receptor-1 associated protein (TRADD) mRNA, 3' end

## DEFINITION

of cds.

## ACCESSION

U41690.1 GI:808914

## VERSION

1.0

## KEYWORDS

TNF receptor; TRADD gene; tumor necrosis factor receptor; tumor necrosis factor receptor 1; tumor necrosis factor receptor-1 associated protein.

## SOURCE

Homo sapiens.

## ORGANISM

Homo sapiens

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;

Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.

REFERENCE 1 (bases 1 to 1435)

AUTHORS Hsu,H., Xiong,J. and Goeddel,D.V.

TITLE The TNF receptor 1-associated protein TRADD signals cell death and NF-kappa B activation (1995)

JOURNAL Cell 81 (4), 495-504 (1995)

MEDLINE 95277836

PUBMED 7758105

FEATURES

source

1..1435

/organism="Homo sapiens"

/db\_xref="taxon:9606"

1..1435

/gene="TRADD"

<1..987

/gene="TRADD"

/standard\_name="TNF receptor-1 associated protein"

/codon\_start=1

/product="tumor necrosis factor receptor type 1 associated protein"

/protein\_id="AA98482.1"

/translation="LAGVGTQAPRRPGRGEMAGONGHEEMVGSAYLVFSSLDKVLSDAHNQQKAVYVRAIQALALABGSPDLQMTKTRSPQILVLRFCGRQCCGF

BASE COUNT 245 a 437 c 494 g 259 t  
 LYPQAFQLRRFVQAEGRATLQRLVEALEENBLTSLAEDLLGLDPMNGIA"  
 ORIGIN

Query Match 100.0%; Score 1435; DB 9; Length 1435;  
 Best Local Similarity 100.0%; Pred. No. 2.3e-220;  
 Matches 1435; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 CTGGGGGCGTGGGAACCCAGGCCCCGCGAGGGGCGGAGAGGATGAGATGGACGCTGG 60  
 DB 1 CTGGGGGCGTGGGAACCCAGGCCCCGCGAGGGGCGGAGAGGATGAGATGGACGCTGG 60  
 QY 61 CAAATGGGCGAGAAAGTGGGTGGGCGAGCCATACCTGTTTGGAGATCCTCGCTGGAC 120  
 DB 61 CAAATGGGCGAGAAAGTGGGTGGGCGAGCCATACCTGTTTGGAGATCCTCGCTGGAC 120  
 QY 121 AAGGTGCTCTGTGGATGCTTACGCGACCCCGACGAGAAAGTGGCGATGACAGGCT 180  
 DB 121 AAGGTGCTCTGTGGATGCTTACGCGACCCCGACGAGAAAGTGGCGATGACAGGCT 180  
 QY 181 CTGAGAGCTGCTTGGCGAGAGGGGGGAGCCCGGAGCGTGGATGCTGAAGATC 240  
 DB 181 CTGAGAGCTGCTTGGCGAGAGGGGGGAGCCCGGAGCGTGGATGCTGAAGATC 240  
 QY 241 CACCGCAGCGACCGCAGCTGATGCTGAGCTGCGATTCTGCGGGCGGACGCTGTGGC 300  
 DB 241 CACCGCAGCGACCGCAGCTGATGCTGAGCTGCGATTCTGCGGGCGGACGCTGTGGC 300  
 QY 301 CGCTTCTCGCGCTTACCGCGAGGGGGCGCTGCGCGCGCTGCGAGAGAGCTTGGCG 360  
 DB 301 CGCTTCTCGCGCTTACCGCGAGGGGGCGCTGCGCGCGCTGCGAGAGAGCTTGGCG 360  
 QY 361 GCGCGGCTCGCGCGCAGCAGCTGCTGCGCGCTGCACTGAGAGCTGCGCGCGCGGCGGAGG 420  
 DB 361 GCGCGGCTCGCGCGCAGCAGCTGCTGCGCGCTGCACTGAGAGCTGCGCGCGCGGCGGAGG 420  
 QY 421 CTGAGACGCTTGTGCGCGAGAGAGCGCTGTTGAGTTGATCTCTAGCCAGAGAGCC 480  
 DB 421 CTGAGACGCTTGTGCGCGAGAGAGCGCTGTTGAGTTGATCTCTAGCCAGAGAGCC 480  
 QY 481 GACCGGCTCGCGGATGAAGAACTGGCTGAGCTGAGAGATGCGCTGCGAAATCTGAAGTGC 540  
 DB 481 GACCGGCTCGCGGATGAAGAACTGGCTGAGCTGAGAGATGCGCTGCGAAATCTGAAGTGC 540  
 QY 541 GCGCTCGGGGGCGCGGGGGTGGAGAGGGAGGTCGTTGGGCCCCCTTGCAGACCCCGGTG 600  
 DB 541 GCGCTCGGGGGCGCGGGGGTGGAGAGGGAGGTCGTTGGGCCCCCTTGCAGACCCCGGTG 600  
 QY 601 CCTCTCTGTGCGAGAGTGAAGCGCGCGCGCGCCAGCCTGCGCAGACTTTTCTGTTT 660  
 DB 601 CCTCTCTGTGCGAGAGTGAAGCGCGCGCGCGCCAGCCTGCGCAGACTTTTCTGTTT 660  
 QY 661 CAGGGTCAAGCTGTAGTGAATCGGCGGCTGAGCTGAAGAGCAACAGAGTTCGCGGCG 720  
 DB 661 CAGGGTCAAGCTGTAGTGAATCGGCGGCTGAGCTGAAGAGCAACAGAGTTCGCGGCG 720  
 QY 721 TCTGTGGGTCTCAATATGCGCGAGGTGGGGGCTTCACTGCGAGCGAGGTGCGGGCGCTG 780  
 DB 721 TCTGTGGGTCTCAATATGCGCGAGGTGGGGGCTTCACTGCGAGCGAGGTGCGGGCGCTG 780  
 QY 781 CCGGACCCGCGCGCTGAGCTCGCTGAGCTTACAGATGAGAGCGGAGAGGACTGTACAGAG 840  
 DB 781 CCGGACCCGCGCGCTGAGCTCGCTGAGCTTACAGATGAGAGCGGAGAGGACTGTACAGAG 840  
 QY 841 GCTTTCAGCTGCTGCGGCGCTTGTGCGAGGCGGAGGGCGCGCGGCGCAGCTGCGAGGC 900  
 DB 841 GCTTTCAGCTGCTGCGGCGCTTGTGCGAGGCGGAGGGCGCGCGGCGCAGCTGCGAGGC 900  
 QY 901 CTGTGTGAGGAGCACTGAGAGAGAGAGAGAGCTTCAACAGCTGCGAGAGAGATTTCTGGGCGCTG 960  
 DB 901 CTGTGTGAGGAGCACTGAGAGAGAGAGAGAGCTTCAACAGCTGCGAGAGAGATTTCTGGGCGCTG 960  
 QY 961 ACCGATCCCAATGCGGCGCTGGCTAGACCAAGGGGTGAGCAGCTTTTGAAGAACTGG 1020

DB 961 ACCGATCCCAATGCGGCGCTGGCTAGACCAAGGGGTGAGCAGCTTTTGAAGAACTGG 1020  
 QY 1021 ATGGCTTAGAGTTCCTTTCGCGGCTATTGCTGAACCCCTGTCCATCCAGAGGACCTTGA 1080  
 DB 1021 ATGGCTTAGAGTTCCTTTCGCGGCTATTGCTGAACCCCTGTCCATCCAGAGGACCTTGA 1080  
 QY 1081 AACTCCATTGGGCTTATCTGCTGGAACCTGTGGGGGCGAGAGTTGATTCCTTCCAGAG 1140  
 DB 1081 AACTCCATTGGGCTTATCTGCTGGAACCTGTGGGGGCGAGAGTTGATTCCTTCCAGAG 1140  
 QY 1141 CCAGACCACTGGGGGTGCATCATTTGGGATTTGCTCAGATACCTTGAATGAGTGGG 1200  
 DB 1141 CCAGACCACTGGGGGTGCATCATTTGGGATTTGCTCAGATACCTTGAATGAGTGGG 1200  
 QY 1201 GTGGGGGGGACTTGTCTTGGAGATCAGCTCAGCTTCTCCATCCAGAGGCGGGCTTA 1260  
 DB 1201 GTGGGGGGGACTTGTCTTGGAGATCAGCTCAGCTTCTCCATCCAGAGGCGGGCTTA 1260  
 QY 1261 CAGCCAGCCCTTACAGTTTACATCATGAGACACCTTATCTTGGTCTTGGAGCTTCA 1320  
 DB 1261 CAGCCAGCCCTTACAGTTTACATCATGAGACACCTTATCTTGGTCTTGGAGCTTCA 1320  
 QY 1321 CTTGGGTCTGCGAGATCTGAGTGAAGTAAACAGGAATCAATCTTGCCTGCCCGCAGC 1380  
 DB 1321 CTTGGGTCTGCGAGATCTGAGTGAAGTAAACAGGAATCAATCTTGCCTGCCCGCAGC 1380  
 QY 1381 TCACACTCAGCGTGGGAGCCCGGAATGTTAAGCAATGATTAAGTAAACAGCG 1435  
 DB 1381 TCACACTCAGCGTGGGAGCCCGGAATGTTAAGCAATGATTAAGTAAACAGCG 1435

## RESULT 3

LOCUS 127061 1441 bp DNA linear PAT 06-FEB-1997  
 DEFINITION Sequence 1 from patent US 5563039.  
 ACCESSION 127061  
 VERSION 127061.1 GI:1817837  
 KEYWORDS

## SOURCE

ORGANISM Unknown.  
 UNCLASSIFIED.

## REFERENCE

1 (bases 1 to 1441)  
 Goeddel, D.V. and Hsu, H.

## TITLES

TMP receptor-associated intracellular signaling proteins and

## JOURNAL

Patent: US 5563039-A 1 08-OCT-1996;

## FEATURES

1..1441  
 Location/Qualifiers

BASE COUNT 251 a 437 c 494 g 259 t  
 ORIGIN /organism="unknown"

Query Match 100.0%; Score 1435; DB 6; Length 1441;  
 Best Local Similarity 100.0%; Pred. No. 2.3e-220;  
 Matches 1435; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 CTGGCGGCGTGGGAACCCAGGCCCCGCGAGGGGCGGAGAGGATGAGATGGCAGCTGG 60  
 DB 1 CTGGCGGCGTGGGAACCCAGGCCCCGCGAGGGGCGGAGAGGATGAGATGGCAGCTGG 60  
 QY 61 CAAATGGGCGAGAAAGTGGGTGGGCGAGCCATACCTGTTTGGAGATCCTCGCTGGAC 120  
 DB 61 CAAATGGGCGAGAAAGTGGGTGGGCGAGCCATACCTGTTTGGAGATCCTCGCTGGAC 120  
 QY 121 AAGGTGCTCTGTGGATGCTTACGCGACCCCGACGAGAAAGTGGCGATGACAGGCT 180  
 DB 121 AAGGTGCTCTGTGGATGCTTACGCGACCCCGACGAGAAAGTGGCGATGACAGGCT 180  
 QY 181 CTGAGAGCTGCTTGGCGAGAGGGGGGAGCCCGGAGCGTGGCTGAGATGCTGAAGATC 240  
 DB 181 CTGAGAGCTGCTTGGCGAGAGGGGGGAGCCCGGAGCGTGGCTGAGATGCTGAAGATC 240

```

QY 241 CACCGGAGGACCCGAGCTGATCTGTGACAGCTGCAATCTGCGGGGCGAGCCCTGTGCG 300
DB 241 CACCGGAGGACCCGAGCTGATCTGTGACAGCTGCAATCTGCGGGGCGAGCCCTGTGCG 300
QY 301 CGCTTCTCCGCGCTTAACCGCGAGGGGGCGCTGCGGCGCGCTGCGAGAGAGCTGTGGG 360
DB 301 CGCTTCTCCGCGCTTAACCGCGAGGGGGCGCTGCGGCGCGCTGCGAGAGAGCTGTGGG 360
QY 361 GCGCGGCTGCGCGAGCACTGATGCGGCTGCACTGGAAGCTGCGCGCGCGCGCGAGCGG 420
DB 361 GCGCGGCTGCGCGAGCACTGATGCGGCTGCACTGGAAGCTGCGCGCGCGCGCGAGCGG 420
QY 421 CTGAGAGCTTTGTGCGGAGCGAGAGCGCGCTTTTGAAGTTGCACTCTAGCCGAGAGCC 480
DB 421 CTGAGAGCTTTGTGCGGAGCGAGAGCGCGCTTTTGAAGTTGCACTCTAGCCGAGAGCC 480
QY 481 GACCGGCTCCGAGATGAAGAACTGCTGAGCTGAGAGATGCGCTGCGAAATCTGAAGTGC 540
DB 481 GACCGGCTCCGAGATGAAGAACTGCTGAGCTGAGAGATGCGCTGCGAAATCTGAAGTGC 540
QY 541 GCGTGGGGGGCGCGGGGCGAGCGGGGAGGTGCTTTCGGGCGCGCTTTCGAGCGCGGGTG 600
DB 541 GCGTGGGGGGCGCGGGGCGAGCGGGGAGGTGCTTTCGGGCGCGCTTTCGAGCGCGGGTG 600
QY 601 CCCTCTCTGTGAGAGTGAAGCGCGCGCGCGCGCGCACTGCGCGAGCTTTTCTGTTTC 660
DB 601 CCCTCTCTGTGAGAGTGAAGCGCGCGCGCGCGCGCACTGCGCGAGCTTTTCTGTTTC 660
QY 661 CAGGGTCAAGCTTGTAGTGAATCGGCGCGCTGAGAGCTGGAAGCAACAGAGTTGCGCGCG 720
DB 661 CAGGGTCAAGCTTGTAGTGAATCGGCGCGCTGAGAGCTGGAAGCAACAGAGTTGCGCGCG 720
QY 721 TCTGTGGGTCTCAATAGGCGCAAGGTGGGGCGTCACTGAGGAGGCGTGGCGGGGCTG 780
DB 721 TCTGTGGGTCTCAATAGGCGCAAGGTGGGGCGTCACTGAGGAGGCGTGGCGGGGCTG 780
QY 781 CCGGACCCGCGCTGAGCTGCTGCGCTTACGAGTACGAGCGCGAGGAGCTGTACGAGCAG 840
DB 781 CCGGACCCGCGCTGAGCTGCTGCGCTTACGAGTACGAGCGCGAGGAGCTGTACGAGCAG 840
QY 841 GCGTTCAGCTGTGCGGCGCTTCTGTGAGAGCGCGCGCGCGCGCGCGCTGCGAGCGG 900
DB 841 GCGTTCAGCTGTGCGGCGCTTCTGTGAGAGCGCGCGCGCGCGCGCGCGCTGCGAGCGG 900
QY 901 CTGTGAGGAGCACTCGAGGAGAGAGAGCTCACGAGCGTGGCAGAGGACTTGTGCGGCGTG 960
DB 901 CTGTGAGGAGCACTCGAGGAGAGAGAGCTCACGAGCGTGGCAGAGGACTTGTGCGGCGTG 960
QY 961 ACCGATCCCAATGCGCGCTGCGCTGAGACCGAGGGGTGCGAGCGCTTTTGGAGAACCTG 1020
DB 961 ACCGATCCCAATGCGCGCTGCGCTGAGACCGAGGGGTGCGAGCGCTTTTGGAGAACCTG 1020
QY 1021 ATGGGCTTGAAGGTTCTTCTGCGGCTATTGCTGAACCCCTGTTCATTCACGGGAGCCCTGA 1080
DB 1021 ATGGGCTTGAAGGTTCTTCTGCGGCTATTGCTGAACCCCTGTTCATTCACGGGAGCCCTGA 1080
QY 1081 AACTCACTTGGCTATCTGCTGAGACCTGCTGGGGGAGAGTTGATGCTTCCCGCAGAG 1140
DB 1081 AACTCACTTGGCTATCTGCTGAGACCTGCTGGGGGAGAGTTGATGCTTCCCGCAGAG 1140
QY 1141 CAGAGACCACTGGGGGTGATCATTTGGGGAATTGCTCTCAGGTACTTTGATAGAGTGTGG 1200
DB 1141 CAGAGACCACTGGGGGTGATCATTTGGGGAATTGCTCTCAGGTACTTTGATAGAGTGTGG 1200
QY 1201 GTGGGGGGGAGCTTGTGAGATCAGCTCACCTTCTCCCATCCAGAAAGCGGGGCTTGA 1260
DB 1201 GTGGGGGGGAGCTTGTGAGATCAGCTCACCTTCTCCCATCCAGAAAGCGGGGCTTGA 1260
QY 1261 CAGCAGAGCCCTTACAGTTTCACTCAGAGAGCACTTGTGATGCTTCTGAGACTTTCAT 1320
DB 1261 CAGCAGAGCCCTTACAGTTTCACTCAGAGAGCACTTGTGATGCTTCTGAGACTTTCAT 1320
QY 1321 CTTGGGTGTGAGATATGCACTGAAGTAAACAGAGAAATCAATCTTGGCTGCGCCCAAG 1380

```

```

DB 1321 CTTGGGTGTGAGATATGCACTGAAGTAAACAGAGAAATCAATCTTGGCTGCGCCCAAGC 1380
QY 1381 TCACACTGAGCTGGAGACCCGAAATGTTAAGCAAGATTAATTAATTAACACGG 1435
DB 1381 TCACACTGAGCTGGAGACCCGAAATGTTAAGCAAGATTAATTAATTAACACGG 1435

RESULT 4
BC004491
LOCUS BC004491 1475 bp mRNA linear PRI 12-JUL-2001
DEFINITION Homo sapiens, clone MGC:11078 IMAGE:3689007, mRNA, complete cds.
ACCESSION BC004491
VERSION BC004491.1 GI:13325366
KEYWORDS MGC.
SOURCE Homo sapiens.
ORGANISM Homo sapiens.
REFERENCE 1 (bases 1 to 1475)
AUTHORS Strausberg R.
TITLE Direct Submission
JOURNAL Submitted (12-MAR-2001) National Institutes of Health, Mammalian
Gene Collection (MGC), Cancer Genomics Office, National Cancer
Institute, 31 Center Drive, Room 11A03, Bethesda, MD 20892-2590,
USA
REMARK NIH-MGC Project URL: http://mgc.nci.nih.gov
COMMENT Contact: MGC help desk
Email: cgabbs-remail.nih.gov
Tissue Procurement: ATCC
CDNA Library Preparation: Rubin Laboratory
DNA Sequencing by: Institute for Systems Biology
http://www.systemsbio.org
contact: amadan@systemsbio.org
Anup Madan, Rachel Dickhoff, Jessica Fahey, Stephanie Ford, Julia
Greene, Mark Ketteman and Anuradha Madan

Clone distribution: MGC clone distribution information can be found
through the I.M.A.G.E. Consortium/LLNL at: http://image.llnl.gov
Series: IRAL Plate: 14 Row: m Column: 21
This clone was selected for full length sequencing because it
passed the following selection criteria: Hexamer frequency ORF
analysis, Similarity but not identity to protein.

FEATURES
source
location/Qualifiers
1..1475
/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone="MGC:11078 IMAGE:3689007"
/tissue_type="Pancreas, adenocarcinoma"
/clone_id="NTH MGC_39"
/lab_host="DH10B-R"
/note="Vector: pOTB7"
67..1005
/codon_start=1
/product="Unknown (protein for MGC:11078)"
/protein_id="AAH04491.1"
/db_xref="GI:13325367"
/translation="MAAGQNGHEBHWGSATLVFESSLDKVVLSNAYHPQKAVAYRA
LQAALASGSGSDVLOMLKTHRSDDPOLIVLRFGRGPGCFLEAYEGALRALQRS
LAAALASHSVPLQLELRAGAERLDALADBERCLSLAQQPRLRDELEALRDLRS
NKKGSGAGGDSVASAPLOPVPVSLSEVAPPPADOTFLQGPVNVNPLSLKD
QUTFARSVGKMRVGSILRGCGALRDPALDSLAVYEREGLEQAFOLLRRVQAE
GRATLRLVLALEBNELTSLABDLCLTTPNGSLA"

BASE COUNT 271 a 442 c 504 g 258 t
ORIGIN
Query Match 99.9%; Score 1433.4; DB 9; Length 1475;
Best Local Similarity 99.9%; Pred. No. 4.1e-220;
Matches 1434; Conservative 0; Mismatches 1; Indels 0; Gaps 0;
QY 1 CTTGGGGGTGTGAGAACCCAGGCGCGCGCGCGAGAGGTGAGATGCGAGCTGGG 60

```

Db 19 CTGGCGGGCGTGGGAACCCAGGCCCCCGCCAGCGGCGCAGAGAGTGAGATGGCACTGGG 78  
 Qy 61 CAAATATGGGACGAGAGAGTGGGTGGGCGAGGCAATACCTGTTTGGAGTCTCGCTGGAC 120  
 Db 79 CAAATATGGGACGAGAGAGTGGGTGGGCGAGGCAATACCTGTTTGGAGTCTCGCTGGAC 138  
 Qy 121 AAGATGATCTGTGTGGATGCTTACGCGCACCCCGCAGAGAGTGGAGTGAACAGGGCT 180  
 Db 139 AAGATGATCTGTGTGGATGCTTACGCGCACCCCGCAGAGAGTGGAGTGAACAGGGCT 198  
 Qy 181 CTGACAGGCTGCTTGGAGAGAGCGGCGGAGCCCGGACGTCGTCGAGAGTGGAGATATC 240  
 Db 199 CTGAGAGCTGCTTGGAGAGAGCGGCGGAGCCCGGACGTCGTCGAGATGCTGAATATC 258  
 Qy 241 CACCGGACGACCGGACGAGTGTGTCAGCTGAGATTTCGCGGGCGGACGCTGTGGC 300  
 Db 259 CACCGGACGACCGGACGAGTGTGTCAGCTGAGATTTCGCGGGCGGACGCTGTGGC 318  
 Qy 301 CGCTTCTCCGCGCTTACCGCGAGGGGGCGCTGGCGCGCGCTGACAGAGAGCTGGCG 360  
 Db 319 CGCTTCTCCGCGCTTACCGCGAGGGGGCGCTGGCGCGCGCTGACAGAGAGCTGGCG 378  
 Qy 361 GCGGCGCTGCGGCGGACGCTGGTCCGCTGCAATGAGAGCTGGCGCGCGCGCGCGGACCG 420  
 Db 379 GCGGCGCTGCGGCGGACGCTGGTCCGCTGCAATGAGAGCTGGCGCGCGCGCGCGGACCG 438  
 Qy 421 CTGAGAGCTTGTGTGGGCGGACGAGAGAGCGCTGTTTGTGATGCACTTACGCCAGACGCC 480  
 Db 439 CTGAGAGCTTGTGTGGGCGGACGAGAGAGCGCTGTTTGTGATGCACTTACGCCAGACGCC 498  
 Qy 481 GACCGGCTCCGCGGATGAAGAACTGGCTGAGCTGAGAGAGTGCCTGCGAAATCTGAATGTC 540  
 Db 499 GACCGGCTCCGCGGATGAAGAACTGGCTGAGCTGAGAGAGTGCCTGCGAAATCTGAATGTC 558  
 Qy 541 GCGTCCGGGGGCGCGGGGTCGCGAGAGGGAGTGCCTTCCGCCCCCTTTCAGCCCCCGGTG 600  
 Db 559 GCGTCCGGGGGCGCGGGGTCGCGAGAGGGAGTGCCTTCCGCCCCCTTTCAGCCCCCGGTG 618  
 Qy 601 CCCTCTCTGTGCGAGGTGAAGCGCGCGCGCGCGCGCACCTGGCCAGACTTTTCTGTC 660  
 Db 619 CCCTCTCTGTGCGAGGTGAAGCGCGCGCGCGCGCGCACCTGGCCAGACTTTTCTGTC 678  
 Qy 661 CAGGAGTCAAGCTGTAGTGAATCGGCGCGCTGAGCTGAAGAGCAACAGAGCTTCCGCGCG 720  
 Db 679 CAGGAGTCAAGCTGTAGTGAATCGGCGCGCTGAGCTGAAGAGCAACAGAGCTTCCGCGCG 728  
 Qy 721 TCTGTGGGTCTCAATATGCGCGAGGTGGGCGCTCACTGCAAGCGAGCTGCCGCGGCTG 780  
 Db 739 TCTGTGGGTCTCAATATGCGCGAGGTGGGCGCTCACTGCAAGCGAGCTGCCGCGGCTG 798  
 Qy 781 CGGAGACCGGCGGCTGAGCTGGTGGCTTACGAGTACGAGGGCGAGGGACTGTACGAGCAG 840  
 Db 799 CGGAGACCGGCGGCTGAGCTGGTGGCTTACGAGTACGAGGGCGAGGGACTGTACGAGCAG 858  
 Qy 841 GCGTTCAGCTGTGCGCGGCTTCTGTCAGAGCGCGAGGGCGCGCGCGCACGCTGCAAGCG 900  
 Db 859 GCGTTCAGCTGTGCGCGGCTTCTGTCAGAGCGCGAGGGCGCGCGCGCACGCTGCAAGCG 918  
 Qy 901 CTGTGTGAGGAGCTGAGAGAGAGAGAGCTCAACGCTGCGAGAGGACTTGTGGGCGCTG 960  
 Db 919 CTGTGTGAGGAGCTGAGAGAGAGAGAGCTCAACGCTGCGAGAGGACTTGTGGGCGCTG 978  
 Qy 961 ACCGATCCCAATGGCGGCGCTGAGTACGAGGGGTGCAACGCTTTTGGAGAACTGG 1020  
 Db 979 ACCGATCCCAATGGCGGCGCTGAGTACGAGGGGTGCAACGCTTTTGGAGAACTGG 1038  
 Qy 1021 ATGGCTTTAGGGTTCCTTCTGCGGCTATGCTGAAACCCCTGTCCATCCAGGGACCTGTA 1080  
 Db 1039 ATGGCTTTAGGGTTCCTTCTGCGGCTATGCTGAAACCCCTGTCCATCCAGGGACCTGTA 1098  
 Qy 1081 AACTCCACTTGGCTTATCTGCTGAGACTGCTGGGGCAGAGTGAATGCTTCCCGAGAG 1140  
 Db 1099 AACTCCACTTGGCTTATCTGCTGAGACTGCTGGGGCAGAGTGAATGCTTCCCGAGAG 1158

Qy 1141 CCAGACCAGTGGGGGTGCAATCATTTGGGGATTTGCTCAGAGTACTTTGATAGAGTGTGG 1200  
 Db 1159 CCAGACCAGTGGGGGTGCAATCATTTGGGGATTTGCTCAGAGTACTTTGATAGAGTGTGG 1218  
 Qy 1201 GTGGGGGGGACCTGTTGGAGATCAGCTTACCTTCTCCATCCAGAAAGCGGGGCTTA 1260  
 Db 1219 GTGGGGGGGACCTGTTGGAGATCAGCTTACCTTCTCCATCCAGAAAGCGGGGCTTA 1278  
 Qy 1261 CAGCCAGCCCTTACATTCATCATGAAGACCTGATCTTGTGGTCTCGACTTCAT 1320  
 Db 1279 CAGCCAGCCCTTACATTCATCATGAAGACCTGATCTTGTGGTCTCGACTTCAT 1338  
 Qy 1321 CCTGGGTGCTGAGATATCTGCAATGAAGTAAACAGAAATCAATCTTGCCTGCCCGACG 1380  
 Db 1339 CTTGGGTGCTGAGATATCTGCAATGAAGTAAACAGAAATCAATCTTGCCTGCCCGACG 1398  
 Qy 1381 TCACACTCAGCGTGGGACCCCGAATGTTAAGCAATGATATAAAGTATTAACACGG 1435  
 Db 1399 TCACACTCAGCGTGGGACCCCGAATGTTAAGCAATGATATAAAGTATTAACACGG 1453

RESULT 5  
 AK090673 1833 bp mRNA linear PRI 15-JUL-2002  
 LOCUS  
 DEFINITION  
 AK090673 Homo sapiens CDNA FLJ3354 fls, clone BRACE2005124, highly similar to TUMOR NECROSIS FACTOR RECEPTOR TYPE 1 ASSOCIATED DEATH DOMAIN PROTEIN.  
 ACCESSION  
 AK090673 GI:21748880  
 VERSION  
 oligo capping; fls (full insert sequence).  
 KEYWORDS  
 Homo sapiens cerebellum CDNA to mRNA, clone\_1lb:BRACE2  
 SOURCE  
 clone:BRACE2005124.  
 ORGANISM  
 Homo sapiens  
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.

REFERENCE  
 AUTHORS  
 Tanigami,A., Fujiwara,T., Shibahara,T., Goto,Y., Hirao,M., Ozaki,K., Sugiyama,T., Irie,R., Otsuki,T., Sato,H., Makamatsu,A., Ishii,S., Yamamoto,J., Isono,Y., Kawai-Hio,Y., Saito,K., Nishikawa,T., Kimura,K., Yamashita,H., Matsuo,K., Nakamura,Y., Sekine,M., Kikuchi,H., Kanda,K., Magatsuna,M., Murakawa,K., Kanehori,K., Takahashi-Fujii,A., Oshima,A., Sugiyama,A., Kawakami,B., Suzuki,Y., Sugano,S., Nagahari,K., Masuno,Y., Nagai,K. and Isogai,T.  
 NEDO human CDNA sequencing project  
 Unpublished  
 TITLE  
 NEDO human CDNA sequencing project  
 JOURNAL  
 REFERENCE  
 2 (bases 1 to 1833)  
 AUTHORS  
 Isogai,T. and Yamamoto,J.  
 TITLE  
 Direct Submission  
 JOURNAL  
 Submitted (04-JUL-2002) Takao Isogai, FLJ Project (HRI Team); 2-6-7 Kazusa-Kametari, Kisarazu, Chiba 292-0812, Japan  
 (E-mail:genom@celeri.co.jp, Tel:81-438-52-3975, Fax:81-438-52-3986)  
 NEDO human CDNA sequencing project supported by Ministry of Economy, Trade and Industry of Japan; CDNA full insert sequencing: Research Association for Biotechnology (RAB) (CDNA library construction: Helix Research Institute (HRI) (supported by Japan Key Technology Center etc.); 5' - & 3' -end one pass sequencing: RAB, HRI, and Biotechnology Center, National Institute of Technology and Evaluation; clone selection for full insert sequencing: HRI and RAB; annotation: HRI and RAB.

FEATURES  
 source  
 1..1833  
 /organism="Homo sapiens"  
 /db\_xref="taxon:9606"  
 /clone="BRACE2005124"  
 /tissue\_type="cerebellum"  
 /clone\_1lb="BRACE2"  
 /note="Cloning vector: pME18sFLJ3"

BASE COUNT 289 a 588 c 580 g 376 t  
 ORIGIN  
 Query Match 86.7%; Score 1243.8; DB 9; Length 1833;

Best Local Similarity 98.7%; Pred. No. 8.9e-190;  
Matches 1254; Conservative 0; Mismatches 17; Indels 0; Gaps 0;

```

QY 165 GGCAGTGAACAGGAGGCTGCGAGGCTGCTTGGCAGAGAGCGCGGAGACCGCGAGCTGT 224
Db 561 GGGACCGCGCGCGCTTGAACCGCGCGCTCCGAGAGAGCGCGGAGCGCGAGCGTGT 620
QY 225 GCAGATGCTGAAGATCAACCGCAGCGACCGCGAGCTGATCTGTGACGTGCGATTCTGCGG 284
Db 621 GCAATGCTGAAGATCAACCGCAGCGACCGCGAGCTGATCTGTGACGTGCGATTCTGCGG 680
QY 285 GCGGAGCTCTGTGAGCGCTTCTCTCGGCGCTTACCGCGAGGCGGCGCTGCGCGCGCT 344
Db 681 GCGGAGCTCTGTGAGCGCTTCTCTCGGCGCTTACCGCGAGGCGGCGCTGCGCGCGCT 740
QY 345 GCAGAGGAGCGCTGCGGCGCGCTGCGCGACGCTGCGGCGCTGCGAATGAGAGCTGCG 404
Db 741 GCAGAGGAGCGCTGCGGCGCGCTGCGCGACGCTGCGGCGCTGCGAATGAGAGCTGCG 800
QY 405 CGCGCGCGCGCGAGCGCTGCGAGCGCTTGTGCGGAGCGAGCGCGCTGTTGATGTTGAT 464
Db 801 CGCGCGCGCGCGAGCGCTGCGAGCGCTTGTGCGGAGCGAGCGCGCTGTTGATGTTGAT 860
QY 465 CCTAGCCCGAGCGCGCGAGCGCGCTCGGAGTAAAGACTGCTGAGCTGAGAGATGCGCT 524
Db 861 CCTAGCCCGAGCGCGCGAGCGCGCTCGGAGTAAAGACTGCTGAGCTGAGAGATGCGCT 920
QY 525 GCGAATCTGAAGTGGGCTGCGGCGCGCGCGGAGTGGCGAGCGGAGGATGCTTGGCGCG 584
Db 921 GCGAATCTGAAGTGGGCTGCGGCGCGCGCGGAGTGGCGAGCGGAGGATGCTTGGCGCG 980
QY 585 CTTCGAGCGCGCGCGGCTCTCTGTGCGAGGTGAAGCGCGCGCGCGCGCGCACTTGC 644
Db 981 CTTCGAGCGCGCGCGGCTCTCTGTGCGAGGTGAAGCGCGCGCGCGCGCGCACTTGC 1040
QY 645 CCAGACTTTTCTGTTCCAGAGGTGAGCTTGAATCGGCGCGCTGAGCTGAGAGACCA 704
Db 1041 CCAGACTTTTCTGTTCCAGAGGTGAGCTTGAATCGGCGCGCTGAGCTGAGAGACCA 1100
QY 705 ACAGAGCTTCTGCGCGCTGCTGAGGTCTCAAAATGGGCGCAAGTGGGCGCTGCACTGAGCG 764
Db 1101 ACAGAGCTTCTGCGCGCTGCTGAGGTCTCAAAATGGGCGCAAGTGGGCGCTGCACTGAGCG 1160
QY 765 AGGCTGCGCGCGCTGCGGAGCGCGCGCTGAGCTGCTGCGCTTGAAGTGAAGCGCGCA 824
Db 1161 AGGCTGCGCGCGCTGCGGAGCGCGCGCTGAGCTGCTGCGCTTGAAGTGAAGCGCGCA 1220
QY 825 GGGACTGTACGAGCAGGCTTTCAGCTGCTGCGGCGCTTCTGTGCGAGCGCGAGGCGCGCG 884
Db 1221 GGGACTGTACGAGCAGGCTTTCAGCTGCTGCGGCGCTTCTGTGCGAGCGCGAGGCGCGCG 1280
QY 885 GCGCAGCGCTGCGCGCTGCTGAGGAGCCTCGAGGAGAAAGAGCTCAGCAGCGCTGCGAGA 944
Db 1281 GCGCAGCGCTGCGCGCTGCTGAGGAGCCTCGAGGAGAAAGAGCTCAGCAGCGCTGCGAGA 1340
QY 945 GGAATTGCTGCGCGCTGCGAGCTTTCAGCTGCTGCGGCGCTTCTGTGCGAGCGCGAGCGCGAG 1004
Db 1341 GGAATTGCTGCGCGCTGCGAGCTTTCAGCTGCTGCGGCGCTTCTGTGCGAGCGCGAGCGCGAG 1400
QY 1005 CTTTGTGAGAACTGAGTGGCTTGAAGGTTCTTCTGCGGCTTCTGTGAGATCCCTGCTGC 1064
Db 1401 CTTTGTGAGAACTGAGTGGCTTGAAGGTTCTTCTGCGGCTTCTGTGAGATCCCTGCTGC 1460
QY 1065 ATCCACGGGAGCGCTGAGAACTCCCACTTGGGCTTATCTGTGAGACCTGCTGGGCGAGATTGA 1124
Db 1461 ATCCACGGGAGCGCTGAGAACTCCCACTTGGGCTTATCTGTGAGACCTGCTGGGCGAGATTGA 1520
QY 1125 TTGCTTCCCGAGAGCGAGACCACTGCGGAGTGCATCATTTGGAGATTCTGCTCAGATAC 1184
Db 1521 TTGCTTCCCGAGAGCGAGACCACTGCGGAGTGCATCATTTGGAGATTCTGCTCAGATAC 1580
QY 1185 TTGATAGATGTGGGTGGGGGAGCTTGTGAGATCAGCTTCACTTCTCCATC 1244

```

```

Db 1581 TTGATAGATGTGGGTGGGGGAGCTTGTGAGATCAGCTTCACTTCTCCATC 1640
QY 1245 CCAGAGCGGCGCTTACAGCCAGCTTACAGTTTCACTGATGAGACCTTGTCTTTG 1304
Db 1641 CCAGAGCGGCGCTTACAGCCAGCTTACAGTTTCACTGATGAGACCTTGTCTTTG 1700
QY 1305 GTGTCTGAGACTTATCTGTGCTGTCTGAGATCTGAGATGAAGTAAACAGGAATCAAT 1364
Db 1701 GTGTCTGAGACTTATCTGTGCTGTCTGAGATCTGAGATGAAGTAAACAGGAATCAAT 1760
QY 1365 CTTGCTGCGCGCGCGAGCTCAGCTCAGCGGTGGAGACCGGAATGTTAAGCAATATATAA 1424
Db 1761 CTTGCTGCGCGCGCGAGCTCAGCTCAGCGGTGGAGACCGGAATGTTAAGCAATATATAA 1820
QY 1425 GTATACACGG 1435
Db 1821 GTATACACGG 1831

```

RESULT 6  
AC074143/c 152953 bp DNA linear PRI 01-MAY-2002  
LOCUS  
DEFINITION Homo sapiens chromosome 16 clone RP11-5A19, complete sequence.  
AC074143  
VERSION AC074143.4 GI:20376965  
KEYWORDS  
HTG.  
SOURCE  
Homo sapiens.  
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.

REFERENCE  
AUTHORS DOE Joint Genome Institute.  
TITLE Sequencing of Human Chromosome 16  
JOURNAL Unpublished  
2 (bases 1 to 152953)  
DOE Joint Genome Institute.  
REFERENCE  
AUTHORS DOE Joint Genome Institute.  
TITLE Direct Submission  
JOURNAL Submitted (15-JUL-2000) Production Sequencing Facility, DOE Joint Genome Institute, 2800 Mitchell Drive, Walnut Creek, CA 94598, USA  
3 (bases 1 to 152953)  
DOE Joint Genome Institute.  
REFERENCE  
AUTHORS Submitted (01-MAY-2002) Production Sequencing Facility, DOE Joint Genome Institute, 2800 Mitchell Drive, Walnut Creek, CA 94598, USA  
On May 1, 2002 this sequence version replaced gi:13786417.

COMMENT  
Sequence Quality Assessment:  
This entry has been annotated with sequence quality estimates computed by the Phrap assembly program. All manually edited bases have been reduced to quality zero. Quality levels above 40 are expected to have less than 1 error in 10,000 bp. Base-by-base quality values are not generally visible from the GenBank flat file format but are available as part of this entry's ASN.1 file.

Sequence Quality Assessment:  
This entry has been annotated with sequence quality estimates computed by the Phrap assembly program. All manually edited bases have been reduced to quality zero. Quality levels above 40 are expected to have less than 1 error in 10,000 bp. Base-by-base quality values are not generally visible from the GenBank flat file format but are available as part of this entry's ASN.1 file.

FEATURES  
source Location/Qualifiers

1..152953  
/organism="Homo sapiens"  
/db\_xref="taxon:9606"  
/chromosome="16"  
/clone="RP11-5A19"  
BASE COUNT 38753 a 35341 c 36249 g 42610 t  
ORIGIN

Query Match 70.1%; Score 1005.8; DB 9; Length 152953;  
 Best Local Similarity 84.2%; Pred. No. 5,7e-152;  
 Matches 1254; Conservative 0; Mismatches 17; Indels 218; Gaps 2;

QY 165 GGCAGTGTACAGGGCTCTGACAGGCTGCTTGGCAGAGACCGCGGAGACCCCGAGCTGCT 224  
 Db 124325 GGGAGCCCGCCCGCCCTGACCGCCCGCTCCGACAGAGCGCGCGGAGCCCGGACGTGCT 124266  
 QY 225 GCAGATGTGTAAGATCCACCGCAGACCGCAGCTGATGTGAGCTGAGGATCTGCGG 284  
 Db 124265 GCAGATGTGTAAGATCCACCGCAGACCGCAGCTGATGTGAGCTGAGGATCTGCGG 124206  
 QY 285 GCGCAGACCCCTGTGCGCGCTTCTCCGCGCTTACCGCGAGGCGGCGCTGCGCGCGCT 344  
 Db 124205 GCGCAGACCCCTGTGCGCGCTTCTCCGCGCTTACCGCGAGGCGGCGCTGCGCGCGCT 124146  
 QY 345 GCAGAGAGAGCTGGCGCGCGCGCTTCCGCGCTTACCGAGCTGCGCTGCGAGCTGCG 404  
 Db 124145 GCAGAGAGAGCTGGCGCGCGCGCTTCCGCGCTTACCGAGCTGCGCTGCGAGCTGCG 124086  
 QY 405 GCGCGGCGCGAGGCGGCTGAGCGCTTGTGCGGAGAGGAGGCTGTTGAGTTCAT 464  
 Db 124085 GCGCGGCGCGAGGCGGCTGAGCGCTTGTGCGGAGAGGAGGCTGTTGAGTTCAT 124026  
 QY 465 CCTAGCCGAGC----- 475  
 Db 124025 CCTAGCCGAGGCTGGGCGCGCGCTAGGCTCAGAGATGGGGTAGGGCGGAGTCCGCGCT 123966  
 QY 476 -----AGCCCGACCGGCTCCGGGAGTAAGAAC 502  
 Db 123965 TAAATCCCCCGTGAAGCGACGCTTCATTCCTCCCTAGCCGAGCCGCTCCGGGATGAAGAAC 123906  
 QY 503 TGGCTGAGCTGGAAGATGCGCTGCGAAATCTGAAGTGGGCTCGGGGCGCCGCGGCTGCG 562  
 Db 123905 TGGCTGAGCTGGAAGATGCGCTGCGAAATCTGAAGTGGGCTCGGGGCGCCGCGGCTGCG 123846  
 QY 563 ACCGGGAGAGTCTGCTGCGCGCCCTTGGAGCGCCCGGTGCGCTCTCTGCGAGGTGAAGC 622  
 Db 123845 ACCGGGAGAGTCTGCTGCGCGCCCTTGGAGCGCCCGGTGCGCTCTCTGCGAGGTGAAGC 123786  
 QY 623 CGCGCGCGCGCGCGCACCTGCGCACTTTTCTGTTCCAGGGTCAAGCTGT----- 674  
 Db 123785 CGCGCGCGCGCGCGCACCTGCGCACTTTTCTGTTCCAGGGTCAAGCTGTAGTGAGG 123726  
 QY 675 ----- 674  
 Db 123725 CGCAGGTGAGAGCGCGGGTCAACAGAGATTGAGTCTGGGAGGGGCAAGGGTGTG 123666  
 QY 675 ----- 674  
 Db 123665 CCGGGTGGGAGTGGGTGGACCGGCGCTGGGGGTCCCGCGCTTACCCCGTTTCCCACT 123606  
 QY 675 -----AGTAATCGGCGCGCTGAGCTGAAGAGCAACAGAGCTTCCGCGCTGTG 726  
 Db 123605 CCTCTGAGTGAATCGGCGCGCTGAGCTGAAGAGCAACAGAGCTTCCGCGCTGTG 123546  
 QY 727 GGTTCAAATGCGCAAGTGGGGCGCTCACTGAGCGAGGCTCCCGGCGCTTCCGGGAC 786  
 Db 123545 GGTTCAAATGCGCAAGTGGGGCGCTCACTGAGCGAGGCTCCCGGCGCTTCCGGGAC 123486  
 QY 787 CCGCGCTGAGCTGCTGAGCTTACGAGTACGAGCGCGAGGAGCTGTACGAGCAGGCTTC 846  
 Db 123485 CCGCGCTGAGCTGCTGAGCTTACGAGTACGAGCGCGAGGAGCTGTACGAGCAGGCTTC 123426  
 QY 847 CAGCTGCTGCGGCGCTTGTGTCAGGCGCGGCGCGCGCTGCAAGCTGCGCTGCTG 906  
 Db 123425 CAGCTGCTGCGGCGCTTGTGTCAGGCGCGGCGCGCGCTGCAAGCTGCGCTGCTG 123366  
 QY 907 GAGGCACTCGAGAGAGAGAGCTCAACAGCTGCGAGAGAGCTTGTGCGGCTGAGTCCGAT 966  
 Db 123365 GAGGCACTCGAGAGAGAGAGCTCAACAGCTGCGAGAGAGCTTGTGCGGCTGAGTCCGAT 123306

QY 967 CCNATGCGGCGCTGCGCTTACGAGCGAGGCTGACGACGCTTTTGGAGAACCTGATGCGC 1026  
 Db 123305 CCNATGCGGCGCTGCGCTTACGAGCGAGGCTGACGACGCTTTTGGAGAACCTGATGCGC 123246  
 QY 1027 TTAGGGTCTTCTTGGCGCTTATGCTGAGACCCCTGTCCATCCAGGAGACCTTGAATCC 1086  
 Db 123245 TTAGGGTCTTCTTGGCGCTTATGCTGAGACCCCTGTCCATCCAGGAGACCTTGAATCC 123186  
 QY 1087 ACTTGGCTTATCTGCTGAGACCTGCTGGGCGAGAGTGTATGCTTCCCGAGAGCCAGAC 1146  
 Db 123185 ACTTGGCTTATCTGCTGAGACCTGCTGGGCGAGAGTGTATGCTTCCCGAGAGCCAGAC 123126  
 QY 1147 CACTGGGGGTGATCATTTGGGGAATTCGCTCAGTAATTTGATGAGTGTGGGCTGCGG 1206  
 Db 123125 CACTGGGGGTGATCATTTGGGGAATTCGCTCAGTAATTTGATGAGTGTGGGCTGCGG 123066  
 QY 1207 GGGACTTGTCTTGGAGATCAGCTCACCTTCTCCATCCAGAGGCGGCTTACAGCCA 1266  
 Db 123065 GGGACTTGTCTTGGAGATCAGCTCACCTTCTCCATCCAGAGGCGGCTTACAGCCA 123006  
 QY 1267 GCCCTTACGTTTCACTCATGAGACGCTTATCTTGGTGTCTGAGACTTCAATCCTGCGG 1326  
 Db 123005 GCCCTTACGTTTCACTCATGAGACGCTTATCTTGGTGTCTGAGACTTCAATCCTGCGG 122946  
 QY 1327 TGTGCGAGATATGCAATGTAAGTAAGTAAGTAAGTAAGTAAGTAAGTAAGTAAGTAAG 1386  
 Db 122945 TGTGCGAGATATGCAATGTAAGTAAGTAAGTAAGTAAGTAAGTAAGTAAGTAAGTAAG 122886  
 QY 1387 TCAAGCTGGGAGCCCGAGTGTTAAGCAATGATATTAAGTAAGTAAGTAAGTAAGTAAG 1435  
 Db 122885 TCAAGCTGGGAGCCCGAGTGTTAAGCAATGATATTAAGTAAGTAAGTAAGTAAGTAAG 122837

RESULT 7  
 AC040160 209574 bp DNA linear PRI 31-MAR-2001  
 LOCUS  
 DEFINITION Homo sapiens chromosome 16 clone CTC-277H1, complete sequence.  
 AC040160  
 AC040160.4 GI:13491214  
 VERSION  
 KEYWORDS  
 SOURCE  
 ORGANISM  
 Homo sapiens.  
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.

REFERENCE  
 1 (bases 1 to 209574)  
 DOE Joint Genome Institute and Stanford Human Genome Center.  
 TITLE  
 Direct Submission  
 JOURNAL  
 2 (bases 1 to 209574)  
 REFERENCE  
 2 (bases 1 to 209574)  
 DOE Joint Genome Institute.  
 TITLE  
 Direct Submission  
 JOURNAL  
 Submitted (31-MAR-2001) DOE Joint Genome Institute, 2800 Mitchell  
 Drive, Walnut Creek, CA 94598, USA  
 On Mar 31, 2001 this sequence version replaced gi:9280711.  
 Draft Sequence Produced by DOE Joint Genome Institute  
 www.jgi.doe.gov  
 www.sngc.stanford.edu  
 Quality: Phrap Quality >=40 99.7% of Sequence;  
 Estimated Total Number of Errors is 0.5.  
 SHGC-11402 G14546.

FEATURES  
 source  
 1..209574  
 /organism="Homo sapiens"  
 /db\_xref="taxon:9606"  
 /chromosome="16"  
 /clone="CTC-277H1"



BASE COUNT 50328 a 55727 c 55548 g 47971 t  
ORIGIN

Query Match 70.1%; Score 1005.8; DB 9; Length 209574;  
Best Local Similarity 84.2%; Pred. No. 5.4e-152;  
Matches 1254; Conservative 0; Mismatches 17; Indels 218; Gaps 2;

```
QY 165 GGCAGTGTACAGGGCTCTGTCAGAGCTGCTTGTGACAGAGCGGGAGCCCGAGCTGCT 224
Db 183123 GGGACCGCGCCGCGCTGACCGCCGCTCCGACAGAGCGGGAGCCCGAGCTGCT 183182
QY 225 GCAATGCTGAAGATTCACCGCAGAGGAGCCCGACGCTGATCTGTCAGCTGCTGCGG 284
Db 183183 GCAAGTGTGAAGATTCACCGCAGAGGAGCCCGACGCTGATCTGTCAGCTGCTGCGG 183242
QY 285 GGGGAGAGCCCTGTGGCGCGCTTCTCTCGCGCCTACCGCAGAGGGAGCGCTGCGCGCGCT 344
Db 183243 GCGGAGCCCTGTGGCGCGCTTCTCTCGCGCCTACCGCAGAGGGAGCGCTGCGCGCGCT 183302
QY 345 GCAGAGAGAGCTGGCGCGCGCGCTGCGCCAGCACTTGATGCGCTGCAACTGGAAGCTGCG 404
Db 183303 GCAGAGAGAGCTGGCGCGCGCGCTGCGCCAGCACTTGATGCGCTGCAACTGGAAGCTGCG 183362
QY 405 CGCCGCGCGCGCGCGCGCTGCGCTTGTGCTGCGCGCAGAGAGCGCTGTTTGAAGTTCAT 464
Db 183363 CGCCGCGCGCGCGCGCGCTGCGCTTGTGCTGCGCGCAGAGAGCGCTGTTTGAAGTTCAT 183422
QY 465 CCTAGCGCGCGC----- 475
Db 183423 CTTAGCCCGCAGAGTGGCGCGCGCGCTAGAGGTAGAGTGGGGATGCCGGAT 183482
QY 476 -----AGCCGACCGCGCTCCGGAATGAAGAAC 502
Db 183483 TAAGTCCCCCGTAGAGCGGACGTTCCATTCCTCCAGCCGACCGGCTCCGGAATGAAGAAC 183542
QY 503 TGGCTGAGCTGAGAGATGCGCTGCGCAATCTGAATGCGCGCTGCGGGGCGCCGGGCTGCGG 562
Db 183543 TGGCTGAGCTGAGAGATGCGCTGCGCAATCTGAATGCGCGCTGCGGGGCGCCGGGCTGCGG 183602
QY 563 AGCGGGAAGGTGCTTGGCGCGCGCGCTTGCAGCGCGCGCGCTGCTGTCGAGAGGTGAAC 622
Db 183603 AGCGGGAAGGTGCTTGGCGCGCGCGCTTGCAGCGCGCGCGCTGCTGTCGAGAGGTGAAC 183662
QY 623 CGCGCGCGCGCGCGCGCGCTGCGCGCGCGCTTGTCTGTCAGAGGTGACCGCTGT----- 674
Db 183663 CGCGCGCGCGCGCGCGCGCTGCGCGCGCGCTTGTCTGTCAGAGGTGACCGCTGT----- 183722
QY 675 ----- 674
Db 183723 CGCAGGTGAGAGCGCGGGGTCAACAGAGAGTGTGAGTCTGGGAGGGGAGGGGCTGTTG 183782
QY 675 ----- 674
Db 183783 CCGGGTGGGGAATGGGTGGAGCCGAGCGCTGGGGGTCCCGCCCTCACCCTTTCCCACT 183842
QY 675 -----AGTGAATTCGCGCGCTGAGCTGAAGAGCAACAAGCTTTCGCGGCTCTGTG 726
Db 183843 CCCTGTCAGTGAATTCGCGCGCTGAGCTGAAGAGCAACAAGCTTTCGCGGCTCTGTG 183902
QY 727 GGTCTCAATTCGCGCAAGGTGGGGCGCTCACTGTCAGAGAGGTGCGGGGCGCTGCGGGAGC 786
Db 183903 GGTCTCAATTCGCGCAAGGTGGGGCGCTCACTGTCAGAGAGGTGCGGGGCGCTGCGGGAGC 183962
QY 787 CGGGGCGCTGGAATCGCTGAGCTTACAGTACGAGCGGAGGGAATGTAACAGAGAGGCTTTC 846
Db 183963 CGGGGCGCTGGAATCGCTGAGCTTACAGTACGAGCGGAGGGAATGTAACAGAGAGGCTTTC 184022
QY 847 CAGCTGCTCGCGGCTTCTGTCAGAGCGCGCGCGCGCACTGTCAGAGCGCTGCTGCTG 906
Db 184023 CAGCTGCTCGCGGCTTCTGTCAGAGCGCGCGCGCGCACTGTCAGAGCGCTGCTGCTG 184082
QY 907 GAGGCACTGAGAGAGACGAGCTCACAGAGCTTGGCAGAGAGCTTGTGGGCTGACCGAT 966
|||
```

```
Db 184083 GAGGCACTGAGAGAGACGAGCTCACAGCTTGGCAGAGAGCTTGTGGGCTGACCGAT 184142
QY 967 CCAATGCGCGCTGCGCTTACAGAGGGGTGACAGCACTTTTGGAGAACCTGATGCGC 1026
Db 184143 CCAATGCGCGCTGCGCTTACAGAGGGGTGACAGCACTTTTGGAGAACCTGATGCGC 184202
QY 1027 TTAGGGTCTTCTTGGCGCTTATGCTGAACCCCTGTGCATTCACAGGGAGCTTGAACCTCC 1086
Db 184203 TTAGGGTCTTCTTGGCGCTTATGCTGAACCCCTGTGCATTCACAGGGAGCTTGAACCTCC 184262
QY 1087 ACTTGGCTTATCTGTGTGAGACCTGCGGGGGCAGAGTGAATGCTTCCCAAGAGCCAGAC 1146
Db 184263 ACTTGGCTTATCTGTGTGAGACCTGCGGGGGCAGAGTGAATGCTTCCCAAGAGCCAGAC 184322
QY 1147 CACTGGGGGTGCATCATTTGGGGAATTCGCTCAGATACCTTGTATGATGATGAGGCTG 1206
Db 184323 CACTGGGGGTGCATCATTTGGGGAATTCGCTCAGATACCTTGTATGATGATGAGGCTG 184382
QY 1207 GGGACTTCTTGGAGATCAAGCTTCACCTTCCCATCCCAAGAGCGGGGCTTACAGCA 1266
Db 184383 GGGACTTCTTGGAGATCAAGCTTCACCTTCCCATCCCAAGAGCGGGGCTTACAGCA 184442
QY 1267 GCGCTTACAGTTTCACTGATGAGAGACCTTGATCTTGGTGTCTGCGCACTTCACTGCGG 1326
Db 184443 GCGCTTACAGTTTCACTGATGAGAGACCTTGATCTTGGTGTCTGCGCACTTCACTGCGG 184502
QY 1327 TGTCTGAGATACCTGAGATGAATGAAGAGAGATCAATCTTCTGCGCGCGCTGACAC 1386
Db 184503 TGTCTGAGATACCTGAGATGAATGAAGAGAGATCAATCTTCTGCGCGCGCTGACAC 184562
QY 1387 TCAGCGTGGAGACCCCGAATGTTAAAGCAATGATTAATTAAGTTAAACAGCG 1435
Db 184563 TCAGCGTGGAGACCCCGAATGTTAAAGCAATGATTAATTAAGTTAAACAGCG 184611
```

RESULT 8  
AC008594/c  
LOCUS 228516 bp DNA linear HTG 21-JUN-2000  
DEFINITION Homo sapiens chromosome 16 clone CTA-277H1, WORKING DRAFT SEQUENCE,  
34 unordered pieces.  
AC008594  
AC008594.6 GI:8575891  
VERSION HTG: HTGS\_PHASE1; HTGS\_DRAFT.  
KEYWORDS Homo sapiens.  
SOURCE Homo sapiens.  
ORGANISM Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.  
REFERENCE 1 (bases 1 to 228516)  
AUTHORS DOE Joint Genome Institute.  
TITLE Sequencing of Human Chromosome 16  
JOURNAL Unpublished  
AUTHORS 2 (bases 1 to 228516)  
REFERENCE DOE Joint Genome Institute.  
TITLE Direct Submission  
JOURNAL Submitted (03-AUG-1999) Production Sequencing Facility, DOE Joint  
Genome Institute, 2800 Mitchell Drive, Walnut Creek, CA 94598, USA  
COMMENT On Jun 21, 2000 this sequence version replaced gi:7689774.  
-----Genome Center  
Center: Joint Genome Institute  
Center Code: JGI  
Web site: <http://www.jgi.doe.gov>  
-----  
Project Information  
Center Project Name: 1258007  
Center clone name: C19788KA\_277H1

Summary Statistics  
Consensus quality: 197894 bases at least Q40  
Consensus quality: 211598 bases at least Q30  
Consensus quality: 214743 bases at least Q20  
Estimated insert size: 212740; agarose-fp estimation  
Estimated insert size: 225216; sum-of-contrigs estimation  
Quality coverage: 6.02 in Q20 bases; agarose-fp estimation



Quality coverage: 5.68 in Q20 bases; sum-of-contigs estimation.

\* NOTE: This is a 'working draft' sequence. It currently consists of 34 contigs. The true order of the pieces is not known and their order in this sequence record is arbitrary. Gaps between the contigs are represented as runs of N, but the exact sizes of the gaps are unknown. This record will be updated with the finished sequence as soon as it is available and the accession number will be preserved.

```
1 1001: contig of 1001 bp in length
1002 1101: gap of unknown length
1102 2101: contig of 1000 bp in length
2102 2201: gap of unknown length
2202 3354: contig of 1153 bp in length
3355 3455: gap of unknown length
3456 4556: contig of 1102 bp in length
4557 4656: gap of unknown length
4657 5682: contig of 1026 bp in length
5683 5783: gap of unknown length
5784 7157: contig of 1375 bp in length
7158 7258: gap of unknown length
7259 8451: contig of 1104 bp in length
8452 8462: gap of unknown length
8463 9465: contig of 1004 bp in length
9466 9565: gap of unknown length
9566 10598: contig of 1033 bp in length
10599 10698: gap of unknown length
10699 11799: contig of 1101 bp in length
11800 11899: gap of unknown length
11900 13061: contig of 1162 bp in length
13062 13161: gap of unknown length
13162 14164: contig of 1003 bp in length
14165 14264: gap of unknown length
14265 16006: contig of 1742 bp in length
16007 16106: gap of unknown length
16107 18072: contig of 1966 bp in length
18073 18172: gap of unknown length
18173 20504: contig of 2332 bp in length
20505 20604: gap of unknown length
20605 24399: contig of 3795 bp in length
24400 24500: gap of unknown length
24501 28638: contig of 4139 bp in length
28639 32855: contig of 4117 bp in length
32856 32955: gap of unknown length
32956 35940: contig of 2985 bp in length
35941 36040: gap of unknown length
36041 39489: contig of 3449 bp in length
39490 39589: gap of unknown length
39590 41280: contig of 1691 bp in length
41281 41380: gap of unknown length
41381 49404: contig of 8024 bp in length
49405 49504: gap of unknown length
49505 53208: contig of 3704 bp in length
53209 53308: gap of unknown length
53309 58096: contig of 4788 bp in length
58097 58196: gap of unknown length
58197 65406: contig of 7210 bp in length
65407 65506: gap of unknown length
65507 73598: contig of 8092 bp in length
73599 73698: gap of unknown length
73699 79746: contig of 6048 bp in length
79747 79846: gap of unknown length
79847 89098: contig of 9252 bp in length
89099 89198: gap of unknown length
89199 98003: contig of 8805 bp in length
98004 98103: gap of unknown length
98104 108051: contig of 9948 bp in length
108052 108151: gap of unknown length
108152 126494: contig of 18343 bp in length
126495 126594: gap of unknown length
126595 149165: contig of 22571 bp in length
149166 149265: gap of unknown length
```

```
* 149266 177969: contig of 28704 bp in length
* 177970 178069: gap of unknown length
* 178070 228516: contig of 50447 bp in length.
FEATURES
source
1..228516
/organism="Homo sapiens"
/db_xref="taxon:9606"
/chromosome="16"
/clone_lib="CtA-277H1"
```

```
BASE COUNT 53125 a 60116 c 58426 g 53336 t 3513 others
ORIGIN
```

```
Query Match 69.0%; Score 990.8; DB 2; Length 228516;
Best local Similarity 84.0%; Pred. No. 1.3e-149;
Matches 1253; Conservative 0; Mismatches 17; Indels 222; Gaps 3;
```

```
QY 165 GGCAGGTGTACAGGAGCTCTGAGAGCTCTTGGCAGAGAGCGGAGAGCCCGACGTGCT 224
Db 168541 GGCAGCGCCCGCGGCTGACCGCCGCTCCGACAGAGCGGAGAGCCCGACGTGCT 168482
QY 225 GCAGATGCTGAAGATCCACCGCAGCAGCCGACGCTGATGTCAGCTGCGATTTCGCG 284
Db 168481 GCAGATGCTGAAGATCCACCGCAGCAGCCGACGCTGATGTCAGCTGCGATTTCGCG 168422
QY 285 GCGGCGAGCCCTGTGGCGGCTTCTCCGCGCTTACCGCGAGGCGGCGCTCGCCCGCGCT 344
Db 168421 GCGGCGAGCCCTGTGGCGGCTTCTCCGCGCTTACCGCGAGGCGGCGCTCGCCCGCGCT 168362
QY 345 GCAGAGAGAGCTTGGCGCGCGCGCTGCGCCAGCACTGCGTCCGCTCAACTGAGAGCTGCG 404
Db 168361 GCAGAGAGAGCTTGGCGCGCGCGCTGCGCCAGCACTGCGTCCGCTCAACTGAGAGCTGCG 168302
QY 405 CGCGCGCGCGCGAGCGGCTGAGCGCTTGTCTGGCGGAGCAGAGAGCGCTGTTGATTGAT 464
Db 168301 CGCGCGCGCGCGAGCGGCTGAGCGCTTGTCTGGCGGAGCAGAGAGCGCTGTTGATTGAT 168242
QY 465 CTTAGCGCCAGC----- 475
Db 168241 CTTAGCGCCAGCAGAGCTGCGCGCGGCTAGAGTGAAGTGGAGTGGCGGAGTCCCGGCT 168182
QY 476 -----AGCCGACCGCGCTCCGGAGTGAAGAAC 502
Db 168181 TAAGTCCCGGTGAGCGGCAAGTTCCATTCCCTAGCCGACCGCGCTCCGGAGTGAAGAAC 168122
QY 503 TGGCTGAGCTGAGAGATGCGCTGCGCAATCTGAATGTCGCGCTCGGCGGCGCGGCTGCGCG 562
Db 168121 TGGCTGAGCTGAGAGATGCGCTGCGCAATCTGAATGTCGCGCTCGGCGGCGCGGCTGCGCG 168062
QY 563 ACGGGAGAGTGGCTTGGCGCGCGCTTGGCGAGCGCGCGCGCTTCTGTGTGAGGTGAAC 622
Db 168061 ACGGGAGAGTGGCTTGGCGCGCGCTTGGCGAGCGCGCGCGCTTCTGTGTGAGGTGAAC 168002
QY 623 CGCGCGCGCGCGCGCACCTGCGCAGACTTTTCTGTCCAGGGTCAAGCTGT----- 674
Db 168001 CGCGCGCGCGCGCGCACCTGCGCAGACTTTTCTGTCCAGGGTCAAGCTGTAGGTGAAG 167942
QY 675 ----- 674
Db 167941 CGCGAGTGAAGAGCGGGGTATCCCGAGGTGAGTCTGGGAGAGGCGGCGCTCGT 167882
QY 675 ----- 674
Db 167881 TGGCGGTGAGTGAAGTGAAGAGCGGCGCTGGGGGTCCCGCGCTCAACCCGTTTCCCC 167822
QY 675 -----AGTGAATCGGCGCTGAGCTGAAGAGCAACAGAGCTTCCGCGCTCT 723
Db 167821 ACTCCCTGAGTGAATGAGCGCGCTGAGCGCTGAAGAGCAAC-ACAAGAGTTCGCGCTCT 167763
QY 724 GTGGGTCTCAAAATGGCGCAAGTGGGCGCTCACTGACAGAGGCTGCGGCGCGCTGCGG 783
Db 167762 GTGGGTCTCAAAATGGCGCAAGTGGGCGCTCACTGACAGAGGCTGCGGCGCGCTGCGG 167703
```

Source	Accession	Gene	Chromosome	Position	Orientation	Strand	Length	GC Content	GC Skew	GC Bias	GC Bias4	GC Bias8	GC Bias16	GC Bias32	GC Bias64	GC Bias128	GC Bias256	GC Bias512	GC Bias1024	GC Bias2048	GC Bias4096	GC Bias8192	GC Bias16384	GC Bias32768	GC Bias65536	GC Bias131072	GC Bias262144	GC Bias524288	GC Bias1048576	GC Bias2097152	GC Bias4194304	GC Bias8388608	GC Bias16777216	GC Bias33554432	GC Bias67108864	GC Bias134217728	GC Bias268435456	GC Bias536870912	GC Bias1073741824	GC Bias2147483648	GC Bias4294967296	GC Bias8589934592	GC Bias17179869184	GC Bias34359738368	GC Bias68719476736	GC Bias137438953472	GC Bias274877906944	GC Bias549755813888	GC Bias1099511627776	GC Bias2199023255552	GC Bias4398046511104	GC Bias8796093022208	GC Bias17592186044416	GC Bias35184372088832	GC Bias70368744177664	GC Bias140737488355328	GC Bias281474976710656	GC Bias562949953421312	GC Bias1125899906842624	GC Bias2251799813685248	GC Bias4503599627370496	GC Bias9007199254740992	GC Bias18014398509481984	GC Bias36028797018963968	GC Bias72057594037927936	GC Bias144115188075855744	GC Bias288230376151711488	GC Bias576460752303422976	GC Bias1152921504606845952	GC Bias2305843009213691904	GC Bias4611686018427383808	GC Bias9223372036854767616	GC Bias18446744073709535328	GC Bias36893488147419070656	GC Bias73786976294838141312	GC Bias147573952589676282624	GC Bias295147905179352565248	GC Bias590295810358704130496	GC Bias1180591620717408260992	GC Bias2361183241434816521984	GC Bias4722366482869633043968	GC Bias9444732965739266087936	GC Bias18889465931478532175808	GC Bias37778931862957064351616	GC Bias75557863725914128703328	GC Bias151115727511828257406656	GC Bias302231455023656514813312	GC Bias604462910047313029226624	GC Bias120892582009462605845248	GC Bias241785164018925211690496	GC Bias483570328037850423380992	GC Bias9671406560757008467601936	GC Bias19342813121514016953203872	GC Bias38685626243028033906407744	GC Bias77371252486056067812809488	GC Bias154742504972112135625618976	GC Bias309485009944224271251237952	GC Bias618970019888448542512475904	GC Bias1237940039776897090852491808	GC Bias2475880079553794181704983616	GC Bias4951760159107588363409967232	GC Bias9903520318215176726819934464	GC Bias1980704063643035445363986928	GC Bias3961408127286070890727973856	GC Bias7922816254572141781455947712	GC Bias15845632509144283563111954424	GC Bias31691265018288567126223908848	GC Bias63382530036577134244438017696	GC Bias126765060073154268488876135392	GC Bias253530120146308536977752227784	GC Bias507060240292617073955504455568	GC Bias1014120480585234147911088911136	GC Bias20282409611704682838221778222272	GC Bias40564819223409365676443554444544	GC Bias81129638446818731272888890888888	GC Bias16225927689363542545777777777777	GC Bias32451855378727085091555555555555	GC Bias64903710757454170181111111111111	GC Bias12980742154890340363222222222222	GC Bias25961484311780680726444444444444	GC Bias51922968623561361452888888888888	GC Bias10384593724722722289057777777777	GC Bias20769187449445445778115555555555	GC Bias41538374898890891556231111111111	GC Bias83076749797781781112462222222222	GC Bias16615349957556356224924444444444	GC Bias33230699915112712448988888888888	GC Bias664613998302254249797777777777	GC Bias13292279966045084995955555555555	GC Bias26584559932090169911911111111111	GC Bias53169119864180339823822222222222	GC Bias1063382397283606796476444444444444	GC Bias21267647945672135929528888888888	GC Bias42535295891344271858057777777777	GC Bias85070591782688543857115555555555	GC Bias17014118356537708714322222222222	GC Bias34028236713075417428444444444444	GC Bias68056473426150834856888888888888	GC Bias13611294682633016913777777777777	GC Bias27222589362660338227555555555555	GC Bias54445178725320676455111111111111	GC Bias10889035745261335311222222222222	GC Bias21778071490522670722444444444444	GC Bias43556143491045341444888888888888	GC Bias871122869820906828897777777777	GC Bias17422473976181136579555555555555	GC Bias34844947952362273115111111111111	GC Bias696898959047245462322222222222	GC Bias13937979184744909246444444444444	GC Bias27875958369489818492888888888888	GC Bias55751916738979
--------	-----------	------	------------	----------	-------------	--------	--------	------------	---------	---------	----------	----------	-----------	-----------	-----------	------------	------------	------------	-------------	-------------	-------------	-------------	--------------	--------------	--------------	---------------	---------------	---------------	----------------	----------------	----------------	----------------	-----------------	-----------------	-----------------	------------------	------------------	------------------	-------------------	-------------------	-------------------	-------------------	--------------------	--------------------	--------------------	---------------------	---------------------	---------------------	----------------------	----------------------	----------------------	----------------------	-----------------------	-----------------------	-----------------------	------------------------	------------------------	------------------------	-------------------------	-------------------------	-------------------------	-------------------------	--------------------------	--------------------------	--------------------------	---------------------------	---------------------------	---------------------------	----------------------------	----------------------------	----------------------------	----------------------------	-----------------------------	-----------------------------	-----------------------------	------------------------------	------------------------------	------------------------------	-------------------------------	-------------------------------	-------------------------------	-------------------------------	--------------------------------	--------------------------------	--------------------------------	---------------------------------	---------------------------------	---------------------------------	---------------------------------	---------------------------------	---------------------------------	----------------------------------	-----------------------------------	-----------------------------------	-----------------------------------	------------------------------------	------------------------------------	------------------------------------	-------------------------------------	-------------------------------------	-------------------------------------	-------------------------------------	-------------------------------------	-------------------------------------	-------------------------------------	--------------------------------------	--------------------------------------	--------------------------------------	---------------------------------------	---------------------------------------	---------------------------------------	--	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---------------------------------------	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---------------------------------------	---	---	---------------------------------------	---	---	-----------------------

```

/organism="Homo sapiens"
/db_xref="taxon:9606"
gene
    63..373
    /gene="TRADD"
exon
    63..373
    /gene="TRADD"
    /number=5
    /usedln=AJ311614:tradd cds
BASE COUNT      92 a 183 c 197 g 107 t
ORIGIN
Query Match      35.7%; Score 512.2; DB 9; Length 579;
Best Local Similarity 98.3%; Pred. No. 1.5e-72;
Matches 528; Conservative 0; Mismatches 8; Indels 1; Gaps 1;

OY 658 TTCAGGGGTCAAGCTGTGATGTAATCGCCGCTGAGCTGGAAGACCAACAGACGTTCCG 717
      |||||
DB 44 TTCCCACTCCCTCGCAGTGAATCGCCGCTGAGCTGGAAGACCAACAGACGTTCCG 103
      |||||
OY 718 CGCTGTGGGTCTTCAATATGCGCAAGGTGGGGGCTTCACATGCGAGCGCTGCCGGGCG 777
      |||||
DB 104 CGCTGTGGGTCTTCAATATGCGCAAGGTGGGGGCTTCACATGCGAGCGCTGCCGGGCG 163
      |||||
OY 778 CTGGCGGACCCGGCGCTGAGCTCGCTGGCTTACAGATTCAGAGCGGAGGAGACTGTACAG 837
      |||||
DB 164 CTGCGGACCCGGCGCTGAGCTCGCTGGCTTACAGATTCAGAGCGGAGGAGACTGTACAG 223
      |||||
OY 838 CAGGCTTTCAGACTGCTGCGCGCTTTCGTGCAAGGCCAGGGCCGCCGCCACGCTGCAG 897
      |||||
DB 224 CAGGCTTTCAGACTGCTGCGCGCTTTCGTGCAAGGCCAGGGCCGCCGCCACGCTGCAG 283
      |||||
OY 898 CGCCTGTGGAGGAGCACTGAGAGAGAACAGACTCACACGACTTGGCAGAGACTTGTGGGC 957
      |||||
DB 284 CGCCTGTGGAGGAGCACTGAGAGAGAACAGACTCACACGACTTGGCAGAGACTTGTGGGC 343
      |||||
OY 958 CTGACCGATCCCAATATGGCGGCGCTGGCTTACAGCAAGGGGTGCAGCCGTTTGGAGAAC 1017
      |||||
DB 344 CTGACCGATCCCAATATGGCGGCGCTGGCTTACAGCAAGGGGTGCAGCCGTTTGGAGAAC 403
      |||||
OY 1018 TGGATGGCTTATAGGTTTCCTTCTGCGGCTATTTGTGTGAACCCCTGTCCATCCAGGGACC 1077
      |||||
DB 404 TGGATGGCTTATAGGTTTCCTTCTGCGGCTATTTGTGTGAACCCCTGTCCATCCAGGGACC 462
      |||||
OY 1078 TGAATCTCACTTGGGCTTATCTGCTGAGACTCTGCTGGGCAAGATTGCTTCCCAAG 1137
      |||||
DB 463 TGAATCTCACTTGGGCTTATCTGCTGAGACTCTGCTGGGCAAGATTGCTTCCCAAG 522
      |||||
OY 1138 GAGCCAGACCACTGGGGGTGATCATCTGGGGATTTCTGCTCAGGACTTGTATGAG 1194
      |||||
DB 523 GAGCCAGACCACTGGGGGTGATCATCTGGGGATTTCTGCTCAGGACTTGTATGAG 579
      |||||

RESULT 10
AC010821/c 208936 bp DNA linear HTG-13-JUL-2000
LOCUS
DEFINITION Homo sapiens clone RP11-3K18, LOW-PASS SEQUENCE SAMPLING.
ACCESSION AC010821
VERSION AC010821.4 GI:9120942
KEYWORDS HTG; HTGS PHASED.
SOURCE
ORGANISM Homo sapiens.
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominiidae; Homo.
            1 (bases 1 to 208936)
            Birren,B., Linton,L., Nusbaum,C. and Lander,E.
            Homo sapiens, clone RP11-3K18
            Homo sapiens, clone RP11-3K18
            Unpublished
            2 (bases 1 to 208936)
REFERENCE
AUTHORS Birren,B., Linton,L., Nusbaum,C., Lander,E., Allen,N., Anderson,M.,
TITLE Baldwin,J., Barna,N., Beckerly,R., Boguslavskiy,L., Boukhalter,B.,
JOURNAL Brown,A., Castle,A., Colangelo,M., Collins,S., Collymore,A.,
REFERENCE Cooke,P., Dekrellano,K., Dewar,K., Domino,M., Donelan,L., Doyle,M.,
AUTHORS Ferreira,P., Fitzhugh,W., Forrest,C., Funke,R., Gage,D.,

```

TITLE  
JOURNAL  
COMMENT

Galagan, J., Gardyna, S., Grant, G., Hagos, B., Heaford, A., Horton, L., Howland, J.C., Johnson, R., Jones, C., Kann, L., Karatas, A., Klein, J., Lohocsky, J., Lieu, C., Locke, K., Macdonald, P., Margulis, N., McEwan, P., McGurk, A., McKernan, C.H., McLaughlin, J., Meldrum, J., Morrow, J., Naylor, J., Norman, C.H., O'Connor, T., O'Donnell, P., Peterson, K., Pollara, V., Riley, R., Roy, A., Santos, R., Severy, P., Stange-Thomann, N., Stojanovic, N., Subramanian, A., Talamas, J., Teafaye, S., Titrrell, A., Vassiliev, H., Vo, A., Wheeler, J., Wu, X., Wyman, D., Ye, W.J., Zimmer, A. and Zody, M.

Submitted (23-SEP-1999) Whitehead Institute/MIT Center for Genome Research, 320 Charles Street, Cambridge, MA 02141, USA  
On Jul 13, 2000 this sequence version replaced gi:8225265.  
All repeats were identified using RepeatMasker:  
Smit, A.F.A. & Green, P. (1996-1997)  
<http://ftp.genome.washington.edu/RM/RepeatMasker.html>

----- Genome Center  
Center: Whitehead Institute/ MIT Center for Genome Research  
Center code: WIBR  
Web site: <http://www-seq.wi.mit.edu>  
Contact: [sequence.submissions@genome.wi.mit.edu](mailto:sequence.submissions@genome.wi.mit.edu)  
----- Project Information  
Center project name: L2764  
Center clone name: 3\_K\_18  
-----

\* NOTE: This record contains 229 individual  
\* sequencing reads that have not been assembled into  
\* contigs. Runs of N are used to separate the reads  
\* and the order in which they appear is completely  
\* arbitrary. Low-pass sequence sampling is useful for  
\* identifying clones that may be gene-rich and allows  
\* overlap relationships among clones to be deduced.  
\* However, it should not be assumed that this clone  
\* will be sequenced to completion. In the event that  
\* the record is updated, the accession number will  
\* be preserved.

1  
\* 698 797: gap of 100 bp in length  
\* 798 1497: contig of 700 bp in length  
\* 1498 1597: gap of 100 bp in length  
\* 1598 2265: contig of 668 bp in length  
\* 2266 2365: gap of 100 bp in length  
\* 2366 3044: contig of 679 bp in length  
\* 3045 3144: gap of 100 bp in length  
\* 3145 3839: contig of 695 bp in length  
\* 3840 3939: gap of 100 bp in length  
\* 3940 4642: contig of 703 bp in length  
\* 4643 4742: gap of 100 bp in length  
\* 4743 5446: contig of 704 bp in length  
\* 5447 5546: gap of 100 bp in length  
\* 5547 6268: contig of 722 bp in length  
\* 6269 6368: gap of 100 bp in length  
\* 6369 7066: contig of 698 bp in length  
\* 7067 7166: gap of 100 bp in length  
\* 7167 7856: contig of 690 bp in length  
\* 7857 7956: gap of 100 bp in length  
\* 7957 8640: contig of 684 bp in length  
\* 8641 8740: gap of 100 bp in length  
\* 8741 9415: contig of 675 bp in length  
\* 9416 9515: gap of 100 bp in length  
\* 9516 10181: contig of 666 bp in length  
\* 10182 10281: gap of 100 bp in length  
\* 10282 10968: contig of 688 bp in length  
\* 10970 11069: gap of 100 bp in length  
\* 11070 11760: contig of 691 bp in length  
\* 11761 11860: gap of 100 bp in length  
\* 11861 12565: contig of 705 bp in length  
\* 12566 12665: gap of 100 bp in length  
\* 12666 13372: contig of 707 bp in length  
\* 13373 13472: gap of 100 bp in length  
\* 13473 14162: contig of 690 bp in length  
\* 14163 14262: gap of 100 bp in length  
\* 14263 14949: contig of 687 bp in length

\* 14950 15049: gap of 100 bp in length  
\* 15050 15750: contig of 701 bp in length  
\* 15751 15850: gap of 100 bp in length  
\* 15851 16524: contig of 674 bp in length  
\* 16525 16624: gap of 100 bp in length  
\* 16625 17329: contig of 705 bp in length  
\* 17330 17429: gap of 100 bp in length  
\* 17430 18125: contig of 696 bp in length  
\* 18126 18225: gap of 100 bp in length  
\* 18226 18911: contig of 686 bp in length  
\* 18912 19011: gap of 100 bp in length  
\* 19012 19698: contig of 687 bp in length  
\* 19699 19798: gap of 100 bp in length  
\* 19799 20498: contig of 700 bp in length  
\* 20499 20598: gap of 100 bp in length  
\* 20599 21313: contig of 715 bp in length  
\* 21314 21413: gap of 100 bp in length  
\* 21414 22130: contig of 717 bp in length  
\* 22131 22230: gap of 100 bp in length  
\* 22231 22917: contig of 687 bp in length  
\* 22918 23017: gap of 100 bp in length  
\* 23018 23713: contig of 696 bp in length  
\* 23714 23813: gap of 100 bp in length  
\* 23814 24511: contig of 698 bp in length  
\* 24512 24611: gap of 100 bp in length  
\* 24612 25296: contig of 685 bp in length  
\* 25297 25396: gap of 100 bp in length  
\* 25397 26074: contig of 678 bp in length  
\* 26075 26174: gap of 100 bp in length  
\* 26175 26796: contig of 622 bp in length  
\* 26797 26896: gap of 100 bp in length  
\* 26897 27596: contig of 700 bp in length  
\* 27597 27696: gap of 100 bp in length  
\* 27697 28400: contig of 704 bp in length  
\* 28401 28500: gap of 100 bp in length  
\* 28501 29220: contig of 720 bp in length  
\* 29221 29320: gap of 100 bp in length  
\* 29321 30032: contig of 712 bp in length  
\* 30033 30132: gap of 100 bp in length  
\* 30133 30825: contig of 693 bp in length  
\* 30826 30925: gap of 100 bp in length  
\* 30926 31627: contig of 702 bp in length  
\* 31628 31727: gap of 100 bp in length  
\* 31728 32408: contig of 681 bp in length  
\* 32409 32508: gap of 100 bp in length  
\* 32509 33192: contig of 684 bp in length  
\* 33193 33292: gap of 100 bp in length  
\* 33293 33968: contig of 676 bp in length  
\* 33969 34068: gap of 100 bp in length  
\* 34069 34744: contig of 676 bp in length  
\* 34745 34844: gap of 100 bp in length  
\* 34845 35526: contig of 682 bp in length  
\* 35527 35626: gap of 100 bp in length  
\* 35627 36322: contig of 696 bp in length  
\* 36323 36422: gap of 100 bp in length  
\* 36423 37118: contig of 696 bp in length  
\* 37119 37218: gap of 100 bp in length  
\* 37219 37946: contig of 728 bp in length  
\* 37947 38046: gap of 100 bp in length  
\* 38047 38736: contig of 690 bp in length  
\* 38737 38836: gap of 100 bp in length  
\* 38837 39531: contig of 695 bp in length  
\* 39532 39631: gap of 100 bp in length  
\* 39632 40318: contig of 687 bp in length  
\* 40319 40418: gap of 100 bp in length  
\* 40419 41109: contig of 691 bp in length  
\* 41110 41209: gap of 100 bp in length  
\* 41210 41900: contig of 691 bp in length  
\* 41901 42000: gap of 100 bp in length  
\* 42001 42619: contig of 619 bp in length  
\* 42620 42719: gap of 100 bp in length  
\* 42720 43400: contig of 681 bp in length  
\* 43401 43500: gap of 100 bp in length

```
* 43501 44116: contig of 616 bp in length
* 44117 44216: gap of 100 bp
* 44217 44916: contig of 700 bp in length
* 44917 45016: gap of 100 bp
* 45017 45713: contig of 697 bp in length
* 45714 45813: gap of 100 bp
* 45814 46497: contig of 684 bp in length
* 46498 46597: gap of 100 bp
* 46598 47293: contig of 696 bp in length
* 47294 47393: gap of 100 bp
* 47394 48098: contig of 705 bp in length
* 48099 48198: gap of 100 bp
* 48199 48892: contig of 694 bp in length
* 48893 48992: gap of 100 bp
* 48993 49666: contig of 674 bp in length
* 49667 49766: gap of 100 bp
* 49767 50453: contig of 687 bp in length
* 50454 50553: gap of 100 bp
* 50554 51229: contig of 676 bp in length
* 51230 51329: gap of 100 bp
* 51330 52017: contig of 688 bp in length
* 52018 52117: gap of 100 bp
* 52118 52803: contig of 686 bp in length
* 52804 52903: gap of 100 bp
* 52904 53610: contig of 707 bp in length
* 53611 53710: gap of 100 bp
* 53711 54406: contig of 696 bp in length
* 54407 54506: gap of 100 bp
* 54507 55214: contig of 708 bp in length
* 55215 55314: gap of 100 bp
* 55315 56029: contig of 715 bp in length
* 56030 56129: gap of 100 bp
* 56130 56834: contig of 705 bp in length
* 56835 56934: gap of 100 bp
* 56935 57626: contig of 692 bp in length
```

Query Match 30.3%; Score 434.4; DB 2; Length 208936;  
Best Local Similarity 94.0%; Pred. No. 1.7e-60;  
Matches 473; Conservative 0; Mismatches 27; Indels 3; Gaps 2;

```
QY 933 CAGCTGGGAGAGGATTTGCTGGGCTGACCCGATCCCATGGGCGCTGGCCCTAGACCG 992
Db 110442 CAGCTTGGAGAGGCTTCTGCGCC- GCGCGATCCAGGGGGCTGGCCCTAGACCG 110385
QY 993 GGGTGCAGCAGCTTTTGGAGAACTGATGCGCTTAGGCTTCTCTGCGGCTATTGCT 1052
Db 110384 GGGTGCAGCAGC- TTGGAGAACTGATGCGCTTAGGCTTCTCTAGGGGAGAAAGCT 110326
QY 1053 GAACCCCTGTCATCCAGCGGACCTGAACTGCACTTGCGCTATGCTGACCTGCTG 1112
Db 110325 GAACCCCTGTCCAACCCGGGGAGCCCTGAACTGCACTTGCGCTATGCTGACCTGCTG 110266
QY 1113 GGGCAGAGTTGATGCTTCCCGAGAGCCAGACCACTGGGGGTGCATCTTGGGATTC 1172
Db 110265 GGGCAGAGTTGATGCTTCCCGAGAGCCAGACCACTGGGGGTGCATCTTGGGATTC 110206
QY 1173 TGCGTCAAGGACTTTGATGAGTGTGGGGGTGGGGGGACTTGCTTTGAGATCAGCTCA 1232
Db 110205 TGCTCAGGACTTTGATGAGTGTGGGGGTGGGGGGACTTGCTTTGAGATCAGCTCA 110146
QY 1233 CCTTCCCATCCAGAGCGGGGCTTAACAGCCAGCCCTTACAGTTTCACTCAATGAAGA 1292
Db 110145 CCTTCCCATCCAGAGCGGGGCTTAACAGCCAGCCCTTACAGTTTCACTCAATGAAGA 110086
QY 1293 CCTTGATCTTTGGTGTCTGAGACTTCACTCGGGTCTGAGATCTGCAATGAAGTAAA 1352
Db 110085 CCTTGATCTTTGGTGTCTGAGACTTCACTCGGGTCTGAGATCTGCAATGAAGTAAA 110026
QY 1353 ACAGGAATCAATCTTGCGTGGCCCCGAGCTCAGACTCAGGGTGGGAGCCCGGAATGTTAAG 1412
Db 110025 ACAGGAATCAATCTTGCGTGGCCCCGAGCTCAGACTCAGGGTGGGAGCCCGGAATGTTAAG 109966
QY 1413 AATGATATAAATATATACACGG 1435
```

Db 109965 AATGATATAAATATATACACGG 109943

```
RESULT 11
LOCUS 127062 1384 bp DNA linear PAT 06-FEB-1997
DEFINITION Sequence 3 from patent US 5563039.
ACCESSION 127062
VERSION 127062.1 GI:1817838
KEYWORDS
SOURCE
ORGANISM Unknown.
REFERENCE 1 (bases 1 to 1384)
AUTHORS Goeddel, D. V. and Hsu, H.
TITLE TNF receptor-associated intracellular signaling proteins and
METHODS methods of use
JOURNAL Patent: US 5563039-A 3 08-Oct-1996;
FEATURES
source location/Qualifiers
BASE COUNT 308 a 345 c 431 g 300 t
ORIGIN
```

Query Match 29.8%; Score 427.6; DB 6; Length 1384;  
Best Local Similarity 78.5%; Pred. No. 4.6e-59;  
Matches 540; Conservative 0; Mismatches 139; Indels 9; Gaps 2;

```
QY 302 GCTTCTCCGCGGCTTACCCGAGGGGGCTGCGCGCGCTGTCAGAGAGCTTGGCGG 361
Db 2 GCTTCTCCAGCTTACCCGAGGGGGCTGCGCGCGCTGTCAGAGAGCTTGGCGGCGG 61
QY 362 CCGGCGTCCGAGAGCTCGGTGCGCTGCAACTGAGACTCGGCGCGCGCGCGAGCGGC 421
Db 62 CCGGCGTCCGAGAGCTCGGTGCGCTGCAACTGAGACTCGGCGCGCGCGCGAGCGGC 121
QY 422 TGGAGCTTTGCTGCGGAGAGAGCGCTGTTGAGTTGATCCTTACCCAGAGCGCG 481
Db 122 TGGAGCTTTGCTGCGGAGAGAGCGCTGTTGAGTTGATCCTTACCCAGAGCGCG 181
QY 482 ACCGCGCTCCGAGATGAAGAACTGCTGAGCTGAGAGATGCGCTGGAATCTGAAGTGC 541
Db 182 ACCGCGCTCCGAGATGAAGAACTGCTGAGCTGAGAGATGCGCTGGAATCTGAAGTGC 241
QY 542 GCTCGGGGGCGGGGAGTGCAGCGGAGAGTGCCTTGCGGCCCTTGAGCGCCGCGTGC 601
Db 242 ACTGCACTGGCCAGGGGTGAGCATACAGTAGCTTCTGACAGTTGCAAGTTCCCGGTTT 301
QY 602 CCTCTGTGCGAGGTGAAGCGCGCGCGCGCCGCACTGCGCAGACTTTCTGTTC 661
Db 302 CCTCTGTGCGAGGTGAAGCGCGCGCGCGCCGCACTGCGCAGACTTTCTGTTC 355
QY 662 AGGTTCAGCTTGTAGTGAATGGCCGCTGAGCTGGAAGACCAACAGACGTTCCGCCCT 721
Db 356 AGGTTCAGCTTGTAGTGAATGGCCGCTGAGCTGGAAGACCAACAGACGTTCCGCCCT 415
QY 722 CTGTGGGTCTAAATGGGCGAAGTGGGGCTCTCACTGCAAGAGCGTCCGGGGCGCTGC 781
Db 416 CGGTGGGTCTAAATGGGCGAAGTGGGGCTCTCACTGCAAGAGCGTCCGGGGCGCTGC 475
QY 782 GGGAGCCGCGCTGAGCTCGCTGAGCTTACAGATGCAAGCGGAGGAGCTTACAGCAGG 841
Db 476 GAGATCTGCGCTTCACTCGCTGAGCTTACAGATGCAAGCGGAGGAGCTTACAGCAGG 535
QY 842 CCTTCAAGCTGTGCGGCGCTTGTGAGGCGGAGGCGCGCGCCAGCTGACGCGCC 901
Db 536 CCTTCAAGCTGTGCGGCGCTTGTGAGGCGGAGGCGCGCGCCAGCTGACGCGCC 595
QY 902 TGGTGAAGCACTCGAGAGAAAGAGCTCACCAGCTGGC--GAGGAGCTTGTGCGGCGC 958
Db 596 TGGTGAAGCACTCGAGAGAAAGAGCTCACCAGCTTGTGAGGAGGAGCTTGTGCGGCGC 655
```

QY 959 TGACCATCCCAATGCGCGCTGACCTTA 986  
 Db 656 AGCGGAGCCGAGTGGCGGCTGACCTTA 683

RESULT 12  
 HSA311615 927 bp DNA linear PRI 16-JUL-2001  
 LOCUS Homo sapiens partial TRAD gene for TNFRSF1A-associated via death  
 DEFINITION domain, exons 3-4.

ACCESSION AJ311615  
 VERSION AJ311615.1 GI:13940271  
 KEYWORDS TNFRSF1A-associated via death domain; TRAD gene.  
 SOURCE human.  
 ORGANISM Homo sapiens  
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.  
 REFERENCE 1 (bases 1 to 927)  
 AUTHORS Scheuerpflug, C.G., Dechant, M., Fellenberg, J., Ewerbeck, V. and  
 Debatin, K.M.  
 TITLE Sequence, genomic organisation, and mutation analysis of the human  
 TRAD gene in childhood B- and T-lineage acute lymphoblastic  
 leukemia and ALPS  
 JOURNAL Unpublished  
 REFERENCE 2 (bases 1 to 927)  
 AUTHORS Scheuerpflug, C.G.  
 TITLE Direct Submission  
 JOURNAL Submitted (05-MAR-2001) Scheuerpflug C.G., Research, University of  
 Heidelberg, Schlieberacher Landstrasse 200a, 69118 Heidelberg,  
 Germany, 69118, GERMANY  
 COMMENT Related sequences: AJ311614-616.  
 FEATURES  
 location/Qualifiers  
 1..927  
 /organism="Homo sapiens"  
 /db\_xref="taxon:9606"  
 gene 308..866  
 /gene="TRAD"  
 exon 308..585  
 /gene="TRAD"  
 /number=3  
 /usedin=AJ311614:tradd\_cds  
 intron 586..667  
 /gene="TRAD"  
 /number=3  
 668..866  
 /gene="TRAD"  
 /number=4  
 /usedin=AJ311614:tradd\_cds  
 BASE COUNT 122 a 293 c 336 g 176 t  
 ORIGIN

Query Match 27.8%; Score 398.4; DB 9; Length 927;  
 Best Local Similarity 80.2%; Pred. No. 2,3e-54;  
 Matches 518; Conservative 0; Mismatches 46; Indels 82; Gaps 1;

QY 165 GGCAGTGTACAGGAGGCTGTGAGGCTGAGAGAGGCGGAGCCCGAGCGTCT 224  
 Db 273 GGAACCCCGCGGCTGTGAGGCTGTGAGGCTGTGAGAGGCGGAGCCCGAGCGTCT 332  
 QY 225 GCAGATCTGAAGATTCACCGAGCGAGCCCGAGCTGATGTGAGAGCTGCGATTCTCGG 284  
 Db 333 GCAGATCTGAAGATTCACCGAGCGAGCCCGAGCTGATGTGAGAGCTGCGATTCTCGG 392  
 QY 285 GCGGAGCCCTGTGAGGCTGTGAGGCTGTGAGGCTGTGAGGAGGCGGCTGTGAGGCTGTGAGGCT 344  
 Db 393 GCGGAGCCCTGTGAGGCTGTGAGGCTGTGAGGCTGTGAGGAGGCGGCTGTGAGGCTGTGAGGCT 452  
 QY 345 GCAGAGGAGGCTGTGAGGCTGTGAGGCTGTGAGGCTGTGAGGCTGTGAGGCTGTGAGGCTGTGAGGCT 404  
 Db 453 GCAGAGGAGGCTGTGAGGCTGTGAGGCTGTGAGGCTGTGAGGCTGTGAGGCTGTGAGGCTGTGAGGCT 512  
 QY 405 CGCGGAGGCTGTGAGGCTGTGAGGCTGTGAGGCTGTGAGGCTGTGAGGCTGTGAGGCTGTGAGGCT 464

Db 513 CGCGGAGGCGGAGCGGCTGTGAGGCTGTGAGGCTGTGAGGCTGTGAGGCTGTGAGGCTGTGAGGCT 572  
 QY 465 CTTAGCCCGAGC----- 475  
 Db 573 CTTAGCCCGAGCGGCTGTGAGGCTGTGAGGCTGTGAGGCTGTGAGGCTGTGAGGCTGTGAGGCT 632  
 QY 476 -----AGCCGAGCGGCTGTGAGGCTGTGAGGCTGTGAGGCTGTGAGGCTGTGAGGCT 502  
 Db 633 TAAGTCCCGGAGCGGAGGCTGTGAGGCTGTGAGGCTGTGAGGCTGTGAGGCTGTGAGGCTGTGAGGCT 682  
 QY 503 TGCGTGTGAGTGTGAGGCTGTGAGGCTGTGAGGCTGTGAGGCTGTGAGGCTGTGAGGCTGTGAGGCT 562  
 Db 693 TGCGTGTGAGTGTGAGGCTGTGAGGCTGTGAGGCTGTGAGGCTGTGAGGCTGTGAGGCTGTGAGGCT 752  
 QY 563 ACGGAGGCTGTGAGGCTGTGAGGCTGTGAGGCTGTGAGGCTGTGAGGCTGTGAGGCTGTGAGGCT 622  
 Db 753 ACGGAGGCTGTGAGGCTGTGAGGCTGTGAGGCTGTGAGGCTGTGAGGCTGTGAGGCTGTGAGGCT 812  
 QY 623 CGCGGCG 682  
 Db 813 CGCGGCG 872  
 QY 683 GCGCGCTGTGAGGCTGTGAGGCTGTGAGGCTGTGAGGCTGTGAGGCTGTGAGGCTGTGAGGCT 728  
 Db 873 CGCAGGTGTGAGGCTGTGAGGCTGTGAGGCTGTGAGGCTGTGAGGCTGTGAGGCTGTGAGGCT 918

RESULT 13  
 AC124713/2 237933 bp DNA linear HTG 23-JUN-2002  
 LOCUS Mus musculus chromosome UNK clone RP24-116N2, WORKING DRAFT  
 DEFINITION  
 AC124713  
 ACCESSION AC124713.2 GI:21539751  
 VERSION AC124713.2  
 KEYWORDS HTG, HTGS, PHASE1, HTGS, DRAFT.  
 SOURCE house mouse.  
 ORGANISM Mus musculus  
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.  
 REFERENCE 1 (bases 1 to 237933)  
 AUTHORS McPherson, J.D. and Waterston, R.H.  
 TITLE The sequence of Mus musculus clone  
 JOURNAL Unpublished  
 REFERENCE 2 (bases 1 to 237933)  
 AUTHORS McPherson, J.D. and Waterston, R.H.  
 TITLE Direct Submission  
 JOURNAL Submitted (15-JUN-2002) Genome Sequencing Center, 4444 Forest Park  
 Parkway, St. Louis, MO 63108, USA  
 3 (bases 1 to 237933)  
 REFERENCE Submitted (23-JUN-2002) Genome Sequencing Center, 4444 Forest Park  
 Parkway, St. Louis, MO 63108, USA  
 On Jun 23, 2002 this sequence version replaced gi:21427971.

COMMENT  
 ----- Genome Center -----  
 Center: Washington University Genome Sequencing Center  
 Center code: WUGSC  
 Web site: http://genome.wustl.edu/gsc/index.shtml  
 Contact: submissions@wustl.edu  
 Project Information  
 Center project name: M\_BB0116N02  
 ----- Summary Statistics -----  
 Sequencing vector: M13; 0%  
 Sequencing vector: plasmid; 100%  
 Chemistry: Dye-terminator Big Dye; 100% of reads  
 Assembly program: Phrap; version 0.990319  
 Consensus quality: 237142 bases at least Q40  
 Consensus quality: 237539 bases at least Q30  
 Consensus quality: 237879 bases at least Q20  
 Insert size: 18800; agarose-fp



Barnhart, C., Edgar, D., Edwards, C.C., Elhaj, C., Escotto, M.,  
 Fall, T., Ferraguto, D., Flagg, N., Ford, J., Foster, P., Frantz, P.,  
 Gabili, A., Gao, J., Garcia, A., Garner, T., Garza, N., Gill, R.,  
 Gorrell, J.H., Guevara, W., Gunaratne, P., Hale, S., Hamilton, K.,  
 Harris, C., Harris, K., Hart, M., Havlak, P., Hayes, A., Hernandez, J.,  
 Hernandez, O., Hodgson, A., Hogues, M., Holloway, C., Hollins, B.,  
 Homs, F., Howard, S., Huber, J., Hui, S., Hume, J., Jackson, L.E.,  
 Jacobson, B., Jia, Y., Johnson, R., Jolivet, S., Joudah, S.,  
 Karlsson, E., Kelly, S., Khan, U., King, L., Korvah, J., Kovar, C.,  
 Kratovic, J., Kureshi, A., Landry, N., Leal, B., Lewis, L.C., Lewis, L.,  
 Li, J., Li, Z., Lichtarge, O., Lien, C., Liu, J., Liu, W., Louie, H.,  
 Lozano, R.J., Lu, X., Lucier, A., Lucier, R., Luna, A., Ma, J.,  
 Maheshwari, M., Mapua, P., Martin, R., Martindale, A., Martinez, E.,  
 Massey, E., Mawhinney, E., McLeod, M.P., Meador, M., Mei, G., Metzger, M.,  
 Miner, G., Miner, Z., Mitchell, T., Mohabhat, K., Morgan, M., Morris, S.,  
 Moser, M., Neal, D., Newton, J., Newton, N., Nguyen, A., Nguyen, N.,  
 Nguyen, N., Nickerson, B., Nwokwenkwo, S., Ogih, M., Okunoda, G.,  
 Oragunye, N., Oyler, R., Pace, A., Payton, B., Peery, J., Perez, L.,  
 Peters, L., Pickens, R., Primus, E., Pu, L.L., Quiles, M., Ren, Y.,  
 Rivers, M., Rojas, A., Rojubokan, I., Rolfe, M., Ruiz, S., Savery, G.,  
 Scherer, S., Scott, G., Shen, H., Shoshitari, N., Sisson, I.,  
 Sodergren, E., Sonaike, T., Sparks, A., Stanley, H., Stone, H.,  
 Sutton, A., Svatek, A., Tabor, P., Tamerisa, A., Tamerisa, K., Tang, H.,  
 Tansey, J., Taylor, C., Taylor, T., Telford, B., Thomas, N., Thomas, S.,  
 Umani, K., Vaequez, L., Vera, Y., Villalón, D., Vinson, R., Wang, Q.,  
 Wang, S., Ward-Moore, S., Warren, R., Washington, C., Wallington, S.,  
 Williams, G., Williamson, A., Wleczky, R., Wooden, S., Worley, K.,  
 Wu, C., Wu, Y., Wu, Y.F., Zhou, J., Zorrilla, S., Nelson, D.,  
 Weinstein, G., and Gibbs, R.  
 Direct Submission  
 Unpublished  
 2 (bases 1 to 180835)  
 Worley, K.C.  
 Direct Submission  
 Submitted (07-MAY-2002) Human Genome Sequencing Center, Department  
 of Molecular and Human Genetics, Baylor College of Medicine, One  
 Baylor Plaza, Houston, TX 77030, USA  
 3 (bases 1 to 180835)  
 Worley, K.C.  
 Direct Submission  
 Submitted (23-JUL-2002) Human Genome Sequencing Center, Department  
 of Molecular and Human Genetics, Baylor College of Medicine, One  
 Baylor Plaza, Houston, TX 77030, USA  
 On Jul 19, 2002 this sequence version replaced gi:20467973.  
 ----- Genome Center  
 Center: Baylor College of Medicine  
 Center code: BCM  
 Web site: <http://www.hgsc.bcm.tmc.edu/>  
 Contact: hgsc-help@bcm.tmc.edu  
 ----- Project Information  
 Project name: GxSP  
 Center project name: CH230-135H12  
 ----- Summary Statistics  
 Sequencing vector: Plasmid  
 Chemistry: Dye-terminator Big Dye 100% of reads  
 Assembly program: Phrap: version 0.990329  
 Consensus quality: 123294 bases at least Q40  
 Consensus quality: 150934 bases at least Q30  
 Consensus quality: 135748 bases at least Q20  
 -----  
 \* NOTE: Estimated insert size may differ from sequence length  
 \* (see [http://www.hgsc.bcm.tmc.edu/docs/genbank\\_draft\\_data.html](http://www.hgsc.bcm.tmc.edu/docs/genbank_draft_data.html)).  
 \* NOTE: This is a 'working draft' sequence. It currently  
 \* consists of 66 contigs. The true order of the pieces  
 \* is not known and their order in this sequence record is  
 \* arbitrary. Gaps between the contigs are represented as  
 \* runs of N, but the exact sizes of the gaps are unknown.  
 \* This record will be updated with the finished sequence  
 \* as soon as it is available and the accession number will  
 \* be prealtered.  
 \* 1 1314: contig of 1314 bp in length  
 \* 1315 1414: gap of unknown length  
 \* 1415 2575: contig of 1161 bp in length  
 \* 2576 2675: gap of unknown length  
 \* 2676 4270: contig of 1595 bp in length  
 \* 4271 4271: gap of unknown length  
 \* 4371 5486: contig of 1116 bp in length  
 \* 5487 5586: gap of unknown length  
 \* 5587 6854: contig of 1268 bp in length  
 \* 6855 6954: gap of unknown length  
 \* 6955 8187: contig of 1233 bp in length  
 \* 8188 8287: gap of unknown length  
 \* 8288 9815: contig of 1528 bp in length  
 \* 9816 9915: gap of unknown length  
 \* 9916 11334: contig of 1419 bp in length  
 \* 11335 11434: gap of unknown length  
 \* 11435 12514: contig of 1080 bp in length  
 \* 12515 12614: gap of unknown length  
 \* 12615 14035: contig of 1421 bp in length  
 \* 14036 14135: gap of unknown length  
 \* 14136 15352: contig of 1217 bp in length  
 \* 15353 15452: gap of unknown length  
 \* 15453 17088: contig of 1636 bp in length  
 \* 17089 17188: gap of unknown length  
 \* 17189 18530: contig of 1342 bp in length  
 \* 18531 18630: gap of unknown length  
 \* 18631 19810: contig of 1180 bp in length  
 \* 19811 19910: gap of unknown length  
 \* 19911 21697: contig of 1787 bp in length  
 \* 21698 21797: gap of unknown length  
 \* 21798 23605: contig of 1808 bp in length  
 \* 23606 23705: gap of unknown length  
 \* 23706 25371: contig of 1666 bp in length  
 \* 25372 25471: gap of unknown length  
 \* 25472 26748: contig of 1277 bp in length  
 \* 26749 26848: gap of unknown length  
 \* 26849 28303: contig of 1435 bp in length  
 \* 28304 28403: gap of unknown length  
 \* 28404 29757: contig of 1354 bp in length  
 \* 29758 29857: gap of unknown length  
 \* 29858 32189: contig of 2332 bp in length  
 \* 32190 32289: gap of unknown length  
 \* 32290 33955: contig of 1666 bp in length  
 \* 33956 34055: gap of unknown length  
 \* 34056 35591: contig of 1536 bp in length  
 \* 35592 35691: gap of unknown length  
 \* 35692 37655: contig of 1664 bp in length  
 \* 37656 37755: gap of unknown length  
 \* 37756 39536: contig of 1781 bp in length  
 \* 39537 39636: gap of unknown length  
 \* 39637 41600: contig of 1964 bp in length  
 \* 41601 41700: gap of unknown length  
 \* 41701 42821: contig of 1131 bp in length  
 \* 42822 42921: gap of unknown length  
 \* 42922 44830: contig of 1909 bp in length  
 \* 44831 44930: gap of unknown length  
 \* 44931 47127: contig of 2197 bp in length  
 \* 47128 47227: gap of unknown length  
 \* 47229 49514: contig of 2287 bp in length  
 \* 49515 49614: gap of unknown length  
 \* 49615 50854: contig of 1240 bp in length  
 \* 50855 52701: contig of 1747 bp in length  
 \* 52702 52801: gap of unknown length  
 \* 52802 54279: contig of 1478 bp in length  
 \* 54280 54379: gap of unknown length  
 \* 54380 55721: contig of 1342 bp in length  
 \* 55722 55821: gap of unknown length  
 \* 55822 57052: contig of 1231 bp in length  
 \* 57053 57152: gap of unknown length  
 \* 57153 59073: contig of 1921 bp in length  
 \* 59074 59173: gap of unknown length  
 \* 59174 62044: contig of 2871 bp in length  
 \* 62045 62144: gap of unknown length  
 \* 62145 64479: contig of 2335 bp in length  
 \* 64480 64579: gap of unknown length



```

* 64580 66834: contig of 2255 bp in length
* 66935 66934: gap of unknown length
* 66935 68756: contig of 1822 bp in length
* 68757 68856: gap of unknown length
* 68857 72006: contig of 3150 bp in length
* 72007 72106: gap of unknown length
* 72107 75196: contig of 3090 bp in length
* 75197 75296: gap of unknown length
* 75297 77560: contig of 2264 bp in length
* 77561 77661: gap of unknown length
* 77661 80355: contig of 2695 bp in length
* 80356 80455: gap of unknown length
* 80456 82230: contig of 1775 bp in length
* 82231 84912: gap of unknown length
* 84913 85012: contig of 2582 bp in length
* 85013 88417: gap of unknown length
* 88418 88517: contig of 3405 bp in length
* 88518 91442: gap of unknown length
* 91443 91543: contig of 2925 bp in length
* 91543 95387: gap of unknown length
* 95388 95487: contig of 3845 bp in length
* 95488 95487: gap of unknown length
* 95488 98904: contig of 3417 bp in length
* 98905 99004: gap of unknown length
* 99005 101628: contig of 2624 bp in length
* 101629 101728: gap of unknown length
* 101729 105532: contig of 3804 bp in length
* 105533 105632: gap of unknown length
* 105633 109920: contig of 4288 bp in length
* 109921 110021: gap of unknown length
* 110021 112467: contig of 2447 bp in length

```

Query Match 19.0%; Score 272; DB 2; Length 180835;

Best Local Similarity 61.3%; Pred. No. 1.7e-34; Indels 230; Gaps 3;

```

Matches 634; Conservative 0; Mismatches 170; Indels 230; Gaps 3;

Qy 177 GGCCTGACAGGCTGCTTGTGACAGAGACGCGGAGCCCGAGCTGCTGACAGTGTCTGAA 236
Db 145385 GGCCTGACAGTGTGCTTGTGACAGAGAGTGGGAGACCCCGAGTGTCTGACAGTGTCTGAA 145444

Qy 237 GATTCACCGCAGGACCCCGAGCTGATCTGTGACGCTGCGATTCTGGCGGCGGACCCCTG 296
Db 145445 GATTCACCTCAGGACCCCTGACGCTGATCTGTGACGCTTGTGCGGCGGCTGTATG 145504

Qy 297 TGGCGGCTTCTCTCGGCGCTACCGGAGGCGGCGTGGCGCGCGCTGCGAGAGAGCCT 356
Db 145505 CGGCGGCTTCTCTCAAGACCTACCGGAGGCGGCGCTGGCGAGCTGCGAGAGTGTCT 145564

Qy 357 GCGCGGCGGCGCTGCGCAGCACTCGGTGCGCTGCACTGAGACTGCGCGCGCGCGCA 416
Db 145565 GCGCGGCGGCGCTTGGCCCAAGAGCACTGGCGTTGCAACTGGAGTTGCGGTGCTGTGCGGA 145624

Qy 417 GCGGCTGACGCTTGTGCTGCGGAGCAAGAGCGCTGTTGAGTTGATCTTACCCCA--- 473
Db 145625 GCGGCTGACGAGTTGCTGCTACCGAGAGCGCTGTTGAGATTATCATCTTACCCAGAA 145684

Qy 474 ----- 473
Db 145685 GGTGCTGTGTCTGTAGTTAGATGGAGATGGTGGGAGATTGGCGCCCAAAAGCCGACATT 145744

Qy 474 -----GCAGCCCGAACCGGCTCCGGATGAAGAACTGGCTGAGCTGAG 516
Db 145745 ACTGTGACCCCTTGGGAGCGCCGACCGGCTCAGGGAGCAAGACTCGGCGAGCTGGAG 145804

Qy 517 GATGCGCTGCGAAATTGAAATGCGGCTCGGGGCGCCGGGCTGCGACGCGGAGTGTCT 576
Db 145805 GATGAGTTCTGCAAACTGACGTGTGACTCAGCTGCGCAGGCTGAGGCAACACAGGTAGCT 145864

Qy 577 TCGGCGCCCTTGAAGCGCCCGGCGCTCTCTGTGCGAGGTTGAAGCGCGCGCGCGCG 636
Db 145865 CAGGAGGTTTGAAGTCTCTGTGTTCTCTCTCGGCGAGGAGAA-----GCCACTGCGG 145918

Qy 637 CCACGTGCCAGACTTCTTCTGTTCCAGGGTCAGC----- 670

```

```

Db 145919 GCGCGCGGCGACAGCTTCTTCTGTCCATGTGCTGATAGGTGAGTTCAACAGACAGAG 145978
Qy 671 ----- 670
Db 145979 CCGGGTACCAAGCAGATGATGTAGACCAGAAATAGGGCTGGGGCCAGGTATCCCGAGCG 146038
Qy 671 -----CT 672
Db 146039 GGTGCGGAGTGTGTGAATCTCCGGGAAACGATGTACTTACTCTTCTCTCTCTCC 146098
Qy 673 GTATGAAATCGGCGCGCTGAGCTTGAAGACCAACAGAGCTTGGCGGCTCTGTGGTCTC 732
Db 146099 GCAGTGAACCGGCGCTGAATCTGTCAAGACCAAGACGATTCGGCGCTCAGTGGGCTTC 146158
Qy 733 AAATGGCGCAAGTGTGGGCGCTCACTGACGAGGCTGCGGGCGCTGCGGAGCCCGGCG 792
Db 146159 AAGTGGCGAGGTAGGGCGCTGCTGCGAGCTGTCTGTCAAGCACTGAGGAGATCTGCG 146218
Qy 793 CTGACTGCTGCGCTTACAGATGACGAGCGCGAGGAGCTGTACGAGCAGGCTTCCAGCTG 852
Db 146219 CTGACTGCTGCGCTTATGAGTATGAGCGTATGAGGCTATATGACAGGCTTCCAGCTG 146278
Qy 853 CTGCGGCGCTTGTGCTGACGCGGAGGCGCGCGCGCGACGCTTGCACGCGCTGTGAGGCA 912
Db 146279 CTGCGCGCTTTCATCAAGACCGAGGCGCGCGGTGCGACACTGCAACGCTGATGAGGCG 146338
Qy 913 CTGAGAGAAACGAGCTACCAAGCGCTGCGAGGAGCTTGTGGGCTGACGATCCCAAT 972
Db 146339 CTGAGAGAAACGAGCTACCAAGCTGATGACAGAAAGATCTGTAGGCTGAGGCGGAGA 146398
Qy 973 GCGCGCTGCGCTTA 986
Db 146399 GCGCGCGAGGCTTA 146412

```

```

RESULT 15
LOCUS AR211525 200 bp DNA linear PAT 20-JUN-2002
DEFINITION Sequence 3 from patent US 639327.
ACCESSION AR211525
VERSION AR211525.1 GI:21514870
KEYWORDS
SOURCE
ORGANISM unknown.
REFERENCE 1 (bases 1 to 200)
AUTHORS Wallach,D., Boldin,M., Goncharov,T. and Golstey,Y.V.
TITLE Modulators of the function of Fas receptors and other proteins
JOURNAL Patent: US 639327-A 3 04-JUN-2002;
FEATURES
Source location/Qualifiers
1..200
BASE COUNT 30 a 68 c 69 g 33 t
ORIGIN

```

Query Match 13.9%; Score 200; DB 6; Length 200;

Best Local Similarity 100.0%; Pred. No. 1.7e-22; Indels 0; Gaps 0;

```

Matches 200; Conservative 0; Mismatches 0; Indels 0;

Qy 622 CCGCGCGCGCGCGCGCACTGCGCAGACTTCTTCTGTTCAGAGGCTGAGCTGTAGTGAAT 681
Db 1 CCGCGCGCGCGCGCGCACTGCGCAGACTTCTTCTGTTCAGAGGCTGAGCTGTAGTGAAT 60

Qy 682 CCGCGCGCTGAGCTGAAGAGCAACAGACGTTGCGGCGCTCTGTGGGTCTCAAAATGGGCG 741
Db 61 CCGCGCGCTGAGCTGAAGAGCAACAGACGTTGCGGCGCTCTGTGGGTCTCAAAATGGGCG 120

Qy 742 AAGGTGGGCGCTGCTCACTGAGGAGAGGCTGCGGCGGCTTGTGCGGAGCCCGCGCTGAGCTCG 801
Db 121 AAGGTGGGCGCTTCACTGAGGAGAGGCTGCGGCGGCTTGTGCGGAGCCCGCGCTGAGCTCG 180

Qy 802 CTGGCTTACGAGTACGAGCG 821

```

Tue Feb 4 11:18:59 2003

us-09-763-748-1.rge

Page 17

Db 181 CTGGCTTACGAGTACGACG 200

Search completed: February 3, 2003, 15:46:41  
Job time : 4531 secs

**THIS PAGE BLANK**

GenCore version 5.1.3  
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: February 3, 2003, 14:21:53 ; Search time 329 Seconds  
(without alignments)  
9822.546 Million cell updates/sec

Title: US-09-763-748-1  
Perfect score: 1435  
Sequence: 1 ctggcgcgcgctgggaaccca.....gataataagatatacagc 1435

Scoring table: IDENTITY\_NUC  
Gapop 10.0, Gapext 1.0

Searched: 2185239 seqs, 1125999159 residues

Total number of hits satisfying chosen parameters: 4370478

Minimum DB seq length: 0  
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%  
Maximum Match 100%

Listing first 45 summaries

Database :

1: N\_Geneseq\_101002.\*  
1: /SID52/gcgdata/geneseq/geneseq-emb1/NA1980.DAT.\*  
2: /SID52/gcgdata/geneseq/geneseq-emb1/NA1981.DAT.\*  
3: /SID52/gcgdata/geneseq/geneseq-emb1/NA1982.DAT.\*  
4: /SID52/gcgdata/geneseq/geneseq-emb1/NA1983.DAT.\*  
5: /SID52/gcgdata/geneseq/geneseq-emb1/NA1984.DAT.\*  
6: /SID52/gcgdata/geneseq/geneseq-emb1/NA1985.DAT.\*  
7: /SID52/gcgdata/geneseq/geneseq-emb1/NA1986.DAT.\*  
8: /SID52/gcgdata/geneseq/geneseq-emb1/NA1987.DAT.\*  
9: /SID52/gcgdata/geneseq/geneseq-emb1/NA1988.DAT.\*  
10: /SID52/gcgdata/geneseq/geneseq-emb1/NA1989.DAT.\*  
11: /SID52/gcgdata/geneseq/geneseq-emb1/NA1990.DAT.\*  
12: /SID52/gcgdata/geneseq/geneseq-emb1/NA1991.DAT.\*  
13: /SID52/gcgdata/geneseq/geneseq-emb1/NA1992.DAT.\*  
14: /SID52/gcgdata/geneseq/geneseq-emb1/NA1993.DAT.\*  
15: /SID52/gcgdata/geneseq/geneseq-emb1/NA1994.DAT.\*  
16: /SID52/gcgdata/geneseq/geneseq-emb1/NA1995.DAT.\*  
17: /SID52/gcgdata/geneseq/geneseq-emb1/NA1996.DAT.\*  
18: /SID52/gcgdata/geneseq/geneseq-emb1/NA1997.DAT.\*  
19: /SID52/gcgdata/geneseq/geneseq-emb1/NA1998.DAT.\*  
20: /SID52/gcgdata/geneseq/geneseq-emb1/NA1999.DAT.\*  
21: /SID52/gcgdata/geneseq/geneseq-emb1/NA2000.DAT.\*  
22: /SID52/gcgdata/geneseq/geneseq-emb1/NA2001A.DAT.\*  
23: /SID52/gcgdata/geneseq/geneseq-emb1/NA2001B.DAT.\*  
24: /SID52/gcgdata/geneseq/geneseq-emb1/NA2002.DAT.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

#### SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	1435	100.0	1435	21	TRADD intracellular
2	1435	100.0	1441	17	Human TRADD intrac
3	1419.2	98.9	1459	22	AAH99549 Human protein enco
4	427.6	29.8	1384	17	AA741463 Mouse TRADD intrac
5	218.4	15.2	649	24	ABO46658 Oligonucleotide fo
6	218.4	15.2	649	24	ABO46659 Oligonucleotide fo
7	198.8	13.9	649	24	ABO46660 Oligonucleotide fo
8	198.8	13.9	649	24	ABO46661 Oligonucleotide fo
9	71.2	5.0	2744	16	AAQ98470 Misp1-containing p

C	10	69.2	4.8	114955	20	AAV53491	Human adenosine A1
	11	64.4	4.5	114955	19	AAV53491	Human adenosine A1
	12	63.8	4.4	799	20	AAV55831	Nucleotide sequenc
	13	63.8	4.4	1926	21	AAV50252	Epstein Barr virus
	14	63.8	4.4	1926	22	AAV50252	EBV tethering prot
	15	63.8	4.4	2860	21	AAV75454	Nucleotide sequenc
	16	63.8	4.4	2580	21	AAV64275	Epstein-Barr virus
	17	63.8	4.4	5452	20	AAV90923	Anti-sense strand
C	18	63.8	4.4	8705	20	AAV23778	Vector pShuttle DN
	19	63.8	4.4	9600	19	AAV21683	Vector plasmid PCM
	20	63.8	4.4	10380	20	AAV22248	Nucleotide sequenc
	21	63.8	4.4	10596	17	AAV51731	Plasmid pCISBON f
	22	63.8	4.4	10596	20	AAV40348	Plasmid pCISBON f
	23	63.8	4.4	10596	20	AAV15650	Nucleotide sequenc
C	24	63.8	4.4	16080	21	AAV59553	DNA clone pCEC C1.
	25	60	4.2	60	24	ABV41584	Human spliced tran
	26	59	4.1	2175	24	AAV97190	Human metalloprote
	27	57	4.0	14855	24	AAV94858	Human DNA sequence
	28	56.6	3.9	3034	24	ABV59321	Nucleotide sequenc
	29	56.2	3.9	2108	24	ABV67774	Oesophagus cancer
	30	55.6	3.9	9551	20	AAV22301	CDNA encoding a hu
	31	55.4	3.9	14800	24	ABV66291	Lung cancer relate
	32	53.4	3.7	1925	20	AAV90924	Epstein Barr Virus
	33	53.2	3.7	6288	22	AAV50590	CDNA encoding nove
	34	53	3.7	1215	19	AAV23483	Pseudomonas Xcps s
	35	53	3.7	1215	19	AAV13894	Pseudomonas alcali
	36	53	3.7	1215	24	AAV22871	Pseudomonas alcali
	37	53	3.7	17612	19	AAV23494	Pseudomonas Xpc, O
	38	53	3.7	17612	21	AAV13905	Pseudomonas alcali
	39	53	3.7	17612	22	AAV30870	Pseudomonas alcali
	40	53	3.7	17612	24	AAV22882	Pseudomonas alcali
	41	52	3.6	1156	24	AAV22051	Human apolipoprote
	42	51.6	3.6	10549	22	AAV15442	Human nervous syst
	43	51.4	3.6	1167	19	AAV75756	Human apolipopro
	44	51.4	3.6	1157	17	AAV06957	Human apolipopro
	45	51.4	3.6	1157	24	ABV95746	Gene #2244 used to

#### ALIGNMENTS

RESULT 1	AAZ93431	AAZ93431 standard; DNA; 1435 BP.
XX	XX	XX
AC	AAZ93431;	
XX	XX	XX
DT	24-JUL-2000	(first entry)
XX	XX	XX
DE	TRADD intracellular protein coding sequence.	
XX	XX	XX
KW	TRADD; TNF; tumour necrosis factor; NF-kappa-B; apoptosis;	
KW	programmed cell death; antisense; inhibition; treatment; therapy;	
KW	septic shock; inflammation; cancer; antiinflammatory; human; ss.	
XX	XX	XX
OS	Homo sapiens.	
XX	XX	XX
FH	Key	Location/Qualifiers
FT	CDS	1..987
FT		/*tag= a
XX	XX	/product= TRADD protein
XX	XX	XX
PN	WO200012527-A1.	
XX	XX	XX
PD	09-MAR-2000.	
XX	XX	XX
PF	25-AUG-1999;	99WO-US19614.
XX	XX	XX
PR	26-AUG-1998;	98US-0143212.
XX	XX	XX
PA	(ISIS-) ISIS PHARM INC.	
XX	XX	XX
PI	Monia BP, Cowsett LM;	

XX MPI: 2000-237846/20.  
DR P-PSDB: AA83156.

XX New antisense compounds that limit the expression of human TRADD  
PT protein, useful in the treatment and diagnosis of cancer, inflammation  
PT and septic shock

XX Example 13: Page 61-64; 85pp; English.

XX The intracellular protein TRADD has been identified as a critical  
CC link between tumour necrosis factor (TNF) receptor binding and  
CC downstream activation of NF-kappa-B. Overexpression of native TRADD  
CC activates NF-kappa-B in the absence of TNF and dominant negative  
CC mutants of TRADD block TNF-induced NF-kappa-B activation. A second  
CC effect of TNF in many cell types is the induction of apoptosis  
CC (programmed cell death). TRADD overexpression has been shown to  
CC mimic TNF induction of apoptosis as well. Data indicates that TRADD  
CC and other downstream effector proteins are the rate limiting step  
CC of TNF action and would therefore serve as the most efficient  
CC targets for inhibition of TNF-induced events. Antisense  
CC oligonucleotides capable of inhibiting TRADD function may therefore  
CC be useful in a number of therapeutic, diagnostic and research  
CC applications. Inhibiting expression of TRADD by contacting human  
CC cells or tissues with the antisense compound may be used to treat a  
CC disease or condition associated with TRADD expression, for example,  
CC septic shock, inflammation, or cancer. TRADD antisense  
CC oligonucleotides of varying inhibitory capabilities are listed in  
CC GENSEQ records AA293438-293517.

XX Sequence 1435 BP; 245 A; 437 C; 494 G; 259 T; 0 other;

XX Query Match 100.0%; Score 1435; DB 21; Length 1435;

XX Best Local Similarity 100.0%; Pred. No. 2.2e-257; Matches 1435; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 CTGGCGGGGCTGGGAAACCCAGGCCCCCGAGCGGCCAGAGAGTGAATGGAGCTGGG 60  
DB 1 CTGGCGGGGCTGGGAAACCCAGGCCCCCGAGCGGCCAGAGAGTGAATGGAGCTGGG 60  
QY 61 CAAATGGGACGAGAGAGTGGGTGGGACGAGTACCTGTTTGTGGAGTCTTGCTGGAC 120  
DB 61 CAAATGGGACGAGAGAGTGGGTGGGACGAGTACCTGTTTGTGGAGTCTTGCTGGAC 120  
QY 121 AAGGTGGTCTGTGGGATGCTTACGGGCAACCCGACAGAGGTGGCAATGACAGGCT 180  
DB 121 AAGGTGGTCTGTGGGATGCTTACGGGCAACCCGACAGAGGTGGCAATGACAGGCT 180  
QY 181 CTGCAAGCTGCTTGGCAGAGAGCGGCGAGCCGGAAGTGTGAGATGCTGAAGATC 240  
DB 181 CTGCAAGCTGCTTGGCAGAGAGCGGCGAGCCGGAAGTGTGAGATGCTGAAGATC 240  
QY 241 CACCGGAGGACCCGAGCTGATGTGCACTGCGATTCTGCGGGGCGGACCCCTGTGCG 300  
DB 241 CACCGGAGGACCCGAGCTGATGTGCACTGCGATTCTGCGGGGCGGACCCCTGTGCG 300  
QY 301 CGCTTCTCCGCGCTTACCGCGAGGGGGGCGTGGCGCGCGCGCTGAGAGAGCCCTGGG 360  
DB 301 CGCTTCTCCGCGCTTACCGCGAGGGGGGCGTGGCGCGCGCGCTGAGAGAGCCCTGGG 360  
QY 361 GCGGCGCTGCCAGACTCGGTGCGCTGCAACTGAGACTGGCGCGCGCGCGAGCGG 420  
DB 361 GCGGCGCTGCCAGACTCGGTGCGCTGCAACTGAGACTGGCGCGCGCGCGAGCGG 420  
QY 421 CTGAGAGCTTTTGTGGCGAGAGAGCGCTGTTTGAAGTTGCACTCTAGCCAGAGCC 480  
DB 421 CTGAGAGCTTTTGTGGCGAGAGAGCGCTGTTTGAAGTTGCACTCTAGCCAGAGCC 480  
QY 481 GACCGGCTCCGGGATGAAGAATGCTGAGCTGAGAGAGTGGCTGGCAATCTGAAGTGC 540  
DB 481 GACCGGCTCCGGGATGAAGAATGCTGAGCTGAGAGAGTGGCTGGCAATCTGAAGTGC 540  
QY 541 GGCTCGGGGGGCCGGGGTGGCGAGCGGGAGAGTGTGCTTGGCCCTTGGAGCCCCCGGTG 600

DB 541 GGCTCGGGGGGCCGGGGTGGCGAGCGGGAGAGTGTGCTTGGCCCTTGGAGCCCCCGGTG 600  
QY 601 CCTCTCTGTGCGAGAGTGAAGCCGCCGCCGCCGCCCACTGCCCCAATCTTTCTGTTTC 660  
DB 601 CCTCTCTGTGCGAGAGTGAAGCCGCCGCCGCCGCCCACTGCCCCAATCTTTCTGTTTC 660  
QY 661 CAGGGTCAAGCCTGTAGTGAATGAGCCGCTGAGCCCTGAAGAGCAACAGACGTTGGCGGC 720  
DB 661 CAGGGTCAAGCCTGTAGTGAATGAGCCGCTGAGCCCTGAAGAGCAACAGACGTTGGCGGC 720  
QY 721 TCTGTGGTCTCAAAATGCGAGAGTGGGGCGCTCACTGCAACGAGGCTGCGGGCGCTG 780  
DB 721 TCTGTGGTCTCAAAATGCGAGAGTGGGGCGCTCACTGCAACGAGGCTGCGGGCGCTG 780  
QY 781 CCGGACCCGGCGCTGAGTCTGCTGAGCTTACAGATACAGGCGGAGAGACTGTACAGAG 840  
DB 781 CCGGACCCGGCGCTGAGTCTGCTGAGCTTACAGATACAGGCGGAGAGACTGTACAGAG 840  
QY 841 GCCTTCAAGCTGTGCGGGCGCTTGTGAGGCGGAGGCGCGCGCGCGACGCTGAGCGC 900  
DB 841 GCCTTCAAGCTGTGCGGGCGCTTGTGAGGCGGAGGCGCGCGCGCGACGCTGAGCGC 900  
QY 901 CTGTGAGAGGCACTGAGAGAGACAGGCTCACAGCTTGGAGAGAGACTTGTGGGCTTG 960  
DB 901 CTGTGAGAGGCACTGAGAGAGACAGGCTCACAGCTTGGAGAGAGACTTGTGGGCTTG 960  
QY 961 ACCGATCCCAATGAGGCGCTGAGCTTACACAGGCGTGAAGCCAGCTTTTGGAGACCTGG 1020  
DB 961 ACCGATCCCAATGAGGCGCTGAGCTTACACAGGCGTGAAGCCAGCTTTTGGAGACCTGG 1020  
QY 1021 AATGACCTTAGGGTCTCTTGTGGGATGCTTGTGAAACCCCTGTCACCGAGGACCTTGA 1080  
DB 1021 AATGACCTTAGGGTCTCTTGTGGGATGCTTGTGAAACCCCTGTCACCGAGGACCTTGA 1080  
QY 1081 AACTCCACTTGGCTTATCTGTGAGACTGTGAGGCGAGAGTGAATGCTTCCCGAGAG 1140  
DB 1081 AACTCCACTTGGCTTATCTGTGAGACTGTGAGGCGAGAGTGAATGCTTCCCGAGAG 1140  
QY 1141 CCAGACCACTGGGGGTGATCATTTGGGGATTTCTGCTTCAAGTACTTTATATGATGTGG 1200  
DB 1141 CCAGACCACTGGGGGTGATCATTTGGGGATTTCTGCTTCAAGTACTTTATATGATGTGG 1200  
QY 1201 GTGGGGGGGACTTGTGTTGGAGATCAGCTTCCCTTCCCATCCAGAGGGGGCTTGA 1260  
DB 1201 GTGGGGGGGACTTGTGTTGGAGATCAGCTTCCCTTCCCATCCAGAGGGGGCTTGA 1260  
QY 1261 CAGCAGCCCTTACAGTTTCACTATGAAGCACTTGAATCTTGTGTCTTGAAGCTTCAAT 1320  
DB 1261 CAGCAGCCCTTACAGTTTCACTATGAAGCACTTGAATCTTGTGTCTTGAAGCTTCAAT 1320  
QY 1321 CCTGGGTGCTGCAATATCTGAGTGAAGTAAACAGGAATCAATCTTGGCTCCGCCAGC 1380  
DB 1321 CCTGGGTGCTGCAATATCTGAGTGAAGTAAACAGGAATCAATCTTGGCTCCGCCAGC 1380  
QY 1381 TCACACTCAGGCTGGAGCCCGGAATGTTAAGCAATGATTAATTAAGTATTAACCGG 1435  
DB 1381 TCACACTCAGGCTGGAGCCCGGAATGTTAAGCAATGATTAATTAAGTATTAACCGG 1435

RESULT 2  
AAT41462  
ID AAT41462 standard; cDNA: 1441 BP.  
XX  
XX AAT41462;  
AC  
XX  
XX  
DT 16-JAN-1997 (first entry)  
XX  
DE Human TRADD intracellular signalling protein cDNA.  
XX Tumour necrosis factor receptor-1 associated death domain protein;  
XX TRADD; signal transduction; cell growth; cell differentiation;  
KW apoptosis; gene therapy; ds.

XX Homo sapiens.  
OS  
XX  
XX Key Location/Qualifiers  
FH CDS 49..987  
FT /tag= a  
FT  
XX  
XX W09630404-A1.  
PN  
XX  
XX 03-OCT-1996.  
PD  
XX  
XX  
PF 15-MAR-1996; 96MO-US03591.  
PR 31-MAR-1995; 95US-0414625.  
XX  
XX (TULAR-) TULARIK INC.  
PA  
XX  
XX Goeddel J, Hsu H;  
PI  
XX  
XX MPI; 1996-455277/45.  
DR P-PSDB; AAW05528.  
XX  
XX  
PT Tumour necrosis factor receptor 1 associated death domain protein  
PT (TRADD) - used to treat and diagnose diseases associated with  
PT undesirable cell growth, migration and/or differentiation  
PS  
XX  
XX Claim 9; Page 31-33; 42pp; English.  
CC A cDNA sequence (AAT41462) codes for human tumour necrosis factor  
CC receptor-1 (TNF-R1) associated death domain protein (TRADD) (AAW05528),  
CC a member of a novel family of intracellular signalling proteins that  
CC modulate cell growth, differentiation and apoptosis. The sequence  
CC was derived from clones isolated from a HUVEC cDNA library using a  
CC probe obtd. from a yeast two-hybrid system that identified proteins  
CC directly interacting with the intracellular region of TNF-R1. The  
CC cDNA can be utilised in recombinant TRADD prodn., as a probe to  
CC detect e.g. the presence of TRADD death domain genes, and in gene  
CC therapy appls.  
XX  
XX  
SQ Sequence 1441 BP; 251 A; 437 C; 494 G; 259 T; 0 other;  
Query Match 100.0%; Score 1435; DB 17; Length 1441;  
Best Local Similarity 100.0%; Pred. No. 2.2e-257;  
Matches 1435; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
QY 1 CTGCGCGGCGTGGGAACCCAGGCCCCCGAGGCGGCCAGAGGTGAGTGGAGCTGGG 60  
DB 1 CTGCGCGGCGTGGGAACCCAGGCCCCCGAGGCGGCCAGAGGTGAGTGGAGCTGGG 60  
QY 61 CAAATAGGACGAAAGTGGGTGGGAGCGCATACCTGTTGTGAGAGTCTGCTGGAGAC 120  
DB 61 CAAATAGGACGAAAGTGGGTGGGAGCGCATACCTGTTGTGAGAGTCTGCTGGAGAC 120  
QY 121 AAGGTGCTCTGTGAGATGCTTACGCGACCCCGAGAGAAAGTGGAGTGTACAGGGCT 180  
DB 121 AAGGTGCTCTGTGAGATGCTTACGCGACCCCGAGAGAAAGTGGAGTGTACAGGGCT 180  
QY 121 AAGGTGCTCTGTGAGATGCTTACGCGACCCCGAGAGAAAGTGGAGTGTACAGGGCT 180  
DB 121 AAGGTGCTCTGTGAGATGCTTACGCGACCCCGAGAGAAAGTGGAGTGTACAGGGCT 180  
QY 181 CTGAGAGCTGCTTGGAGAGAGCGGCGGAGCCCGAGACGTGTGAGATGCTAAAGATC 240  
DB 181 CTGAGAGCTGCTTGGAGAGAGCGGCGGAGCCCGAGACGTGTGAGATGCTAAAGATC 240  
QY 181 CTGAGAGCTGCTTGGAGAGAGCGGCGGAGCCCGAGACGTGTGAGATGCTAAAGATC 240  
DB 181 CTGAGAGCTGCTTGGAGAGAGCGGCGGAGCCCGAGACGTGTGAGATGCTAAAGATC 240  
QY 241 CACCGCAGCGACCCGAGCTGATGCTGAGCTGAGATTTCTGCGGGCGGAGACCTTGTGGC 300  
DB 241 CACCGCAGCGACCCGAGCTGATGCTGAGCTGAGATTTCTGCGGGCGGAGACCTTGTGGC 300  
QY 301 CGCTTCTCCGCGCTTACCGGAGAGGCGCTGTGCGCCGCGCTGCAAGAGAGCTTGGCG 360  
DB 301 CGCTTCTCCGCGCTTACCGGAGAGGCGCTGTGCGCCGCGCTGCAAGAGAGCTTGGCG 360  
QY 361 GCCCGCTCGCCGAGCACTGGTGCCTGTGCACTGAGAGCTGCCGCGGCGCGAGCGG 420  
DB 361 GCCCGCTCGCCGAGCACTGGTGCCTGTGCACTGAGAGCTGCCGCGGCGCGAGCGG 420

QY 421 CTGAGACGCTTTGCTGGCGGACGAGAGCGCTGTTGAGATTGATCTTACGCCAGACCCC 480  
DB 421 CTGAGACGCTTTGCTGGCGGACGAGAGCGCTGTTGAGATTGATCTTACGCCAGACCCC 480  
QY 481 GACCGGCTCCGGGATGAAGAACTGGCTGAGCTGAGAGATGCGCTGCGAAATCTGAAGTGC 540  
DB 481 GACCGGCTCCGGGATGAAGAACTGGCTGAGCTGAGAGATGCGCTGCGAAATCTGAAGTGC 540  
QY 541 GGCTCGGGGGCCCGGGGGTGGCGAGCGGGGAGGTGCTTCCGCCCCCTTGCAGCCCCGGTG 600  
DB 541 GGCTCGGGGGCCCGGGGGTGGCGAGCGGGGAGGTGCTTCCGCCCCCTTGCAGCCCCGGTG 600  
QY 601 CCTCTCTGTGCGAGGTGAAGCGCGCGCGCGCGCCAGCCTTCCAGACTTTTCTGTTTC 660  
DB 601 CCTCTCTGTGCGAGGTGAAGCGCGCGCGCGCGCCAGCCTTCCAGACTTTTCTGTTTC 660  
QY 661 CAGGCTCAGCTTGTAGTGAATGCGCGCTGAGCTGAGAGCCAGACAGACCTTTCGCGGC 720  
DB 661 CAGGCTCAGCTTGTAGTGAATGCGCGCTGAGCTGAGAGCCAGACAGACCTTTCGCGGC 720  
QY 721 TCTGTGGGTCTCAAAATGAGCGCAAGTGGGGCGCTCACTGCGAGGAGCTGCGGGCGCTG 780  
DB 721 TCTGTGGGTCTCAAAATGAGCGCAAGTGGGGCGCTCACTGCGAGGAGCTGCGGGCGCTG 780  
QY 781 CCGAGACCGGCGCTGAGACTGCTGAGCTTACAGATACGAGCGCGAGGAGCTGTACGAGCAG 840  
DB 781 CCGAGACCGGCGCTGAGACTGCTGAGCTTACAGATACGAGCGCGAGGAGCTGTACGAGCAG 840  
QY 841 GCCTTCCAGCTGCTGCGGGCGCTTGTGCGAGCGCGAGGGCGCGCGCCAGCTGCGAGCGC 900  
DB 841 GCCTTCCAGCTGCTGCGGGCGCTTGTGCGAGCGCGAGGGCGCGCGCCAGCTGCGAGCGC 900  
QY 901 CTGTGGAGGCACTCGAGAGAAAGAGCTCACAGCGCTGCGAGAGACTTGTGCGGGCTG 960  
DB 901 CTGTGGAGGCACTCGAGAGAAAGAGCTCACAGCGCTGCGAGAGACTTGTGCGGGCTG 960  
QY 961 ACCGATCCCAATGCGCGCTGAGACCGAGGGGTGCAAGCAGCTTTTGGAGAACTGG 1020  
DB 961 ACCGATCCCAATGCGCGCTGAGACCGAGGGGTGCAAGCAGCTTTTGGAGAACTGG 1020  
QY 1021 ATGGCTTAAAGTTCCTTCTGCGGCTATTGCTGAACCCCTGTCCATCCAGGGACCTGA 1080  
DB 1021 ATGGCTTAAAGTTCCTTCTGCGGCTATTGCTGAACCCCTGTCCATCCAGGGACCTGA 1080  
QY 1081 AACTCCACTTGGGCTTATCTGCTGAGACCTGCTGGGAGAGATGATTTGCCCTTCCCAAGAG 1140  
DB 1081 AACTCCACTTGGGCTTATCTGCTGAGACCTGCTGGGAGAGATGATTTGCCCTTCCCAAGAG 1140  
QY 1141 CCAGACCACTGGGGGTGCAATCTGGGGGATTTGCTCCAGAGTACTTGTATGAGTGGG 1200  
DB 1141 CCAGACCACTGGGGGTGCAATCTGGGGGATTTGCTCCAGAGTACTTGTATGAGTGGG 1200  
QY 1201 GTGGGGGGGACTTGTGAGATCAGCTTCCCTTCCATCCAGAAAGCGGGGCTTA 1260  
DB 1201 GTGGGGGGGACTTGTGAGATCAGCTTCCCTTCCATCCAGAAAGCGGGGCTTA 1260  
QY 1261 CAGCGAGCCCTTACAGTTTACATGAGAGCACTTGAATTTGTGGTCTCTGAGACTTCAT 1320  
DB 1261 CAGCGAGCCCTTACAGTTTACATGAGAGCACTTGAATTTGTGGTCTCTGAGACTTCAT 1320  
QY 1321 CTTGGGTGCTGAGATTAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAG 1380  
DB 1321 CTTGGGTGCTGAGATTAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAG 1380  
QY 1381 TCACTACTCAGCTGGGAGACCCGAGTGTAAAGCAATGTAATTAAGTGAAGTGAAG 1435  
DB 1381 TCACTACTCAGCTGGGAGACCCGAGTGTAAAGCAATGTAATTAAGTGAAGTGAAG 1435  
RESULT 3  
AAH9549  
ID AAH9549 standard; cDNA; 1459 BP.  
XX

AC AAH99549;  
XX 16-OCT-2001 (first entry)  
DT  
XX  
XX Human protein encoding cDNA sequence SEQ ID NO:384.  
XX  
XX Human; cancer; ulcer; HIV infection; human immunodeficiency virus;  
KW antiinflammatory; antirheumatic; antiallergic; immunosuppressive;  
KW antibacterial; endocrine; cardiac; central nervous system; vincicide;  
XX anti-HIV; fungicide; antimutagen; cardiovascular; antianaemic; anaemia;  
KW antilegionella; haemostatic; vulnery; antitumor; osteoporosis; eczema;  
KW dermatological; antiallergic; antidiabetic; cytosolic;  
KW neuroprotective; antidepressant; nootropic; antiparkinsonian; infection;  
KW immunostimulant; gene therapy; antisense therapy; vaccine; inflammation;  
KW antianaphylactic; rheumatoid arthritis; septic shock; pancreatitis;  
KW cardiac dysfunction; neuropathology; cardiac anaphylaxis; autoimmunity;  
KW genetic disease; haematopoietic disorder; platelet disorder; asthma;  
KW thrombocytopenia; osteoporosis; severe combined immunodeficiency;  
KW allergic rhinitis; diabetes; multiple sclerosis; depression;  
KW Alzheimer's disease; Parkinson's disease; neurodegenerative disorder;  
KW neurological disorder; ss.  
XX  
XX Homo sapiens.  
OS  
XX WO200153455-A2.  
PN  
XX 26-JUN-2001.  
PD  
XX 22-DEC-2000; 2000WO-US35017.  
PF  
XX 23-DEC-1999; 99US-0471275.  
PR 21-JAN-2000; 2000US-0488725.  
PR 25-APR-2000; 2000US-0552317.  
XX  
XX (HYSE-) HYSEQ INC.  
PA  
XX Tang YT, Liu C, Dymanc RT;  
PI  
XX MPI; 2001-457603/49.  
DR P-PSDB; AAM25508.  
XX  
PT Isolated human polynucleotides encoding polypeptides, useful for the  
PT treatment and diagnosis of e.g. cancer, ulcers and HIV infection -  
XX  
XX Claim 1; Page 486-487; 1217pp; English.  
PS  
XX AAH99166 to AAH99904 encode the human proteins given in AAM25225 to  
CC AAM25963. The proteins can have activities based on the tissues and  
CC cells they are expressed in, such as: antiinflammatory; antirheumatic;  
CC antiallergic; immunosuppressive; antibacterial; endocrine; cardiac;  
CC central nervous system; vincicide; anti-HIV; fungicide; antimutagen;  
CC cardiovascular; antianaemic; antilegionella; haemostatic; vulnery;  
CC antitumor; osteoporosis; dermatological; antiallergic; antidiabetic;  
CC antiparkinsonian; and immunostimulant. The proteins and polynucleotides  
CC encoding them can be used in gene therapy, antisense therapy and vaccine  
CC production. The proteins and polynucleotides are useful for screening for  
CC agonists or antagonists of a protein and for the treatment and diagnosis  
CC of disorders associated with the activity of a protein e.g. inflammation,  
CC rheumatoid arthritis, septic shock, pancreatitis, cardiac dysfunction,  
CC neuropathology, cardiac anaphylaxis, viral, bacterial, HIV and fungal  
CC infections, autoimmunity, genetic diseases, haematopoietic disorders,  
CC anaemia, platelet disorders, thrombocytopenia, wounds, burns, ulcers,  
CC osteoporosis, severe combined immunodeficiency, eczema, allergic  
CC rhinitis, asthma, diabetes, cancer, multiple sclerosis, depression,  
CC Alzheimer's disease, Parkinson's disease, neurodegenerative and  
CC neurological disorders.  
XX  
XX Sequence 1459 BP; 258 A; 437 C; 498 G; 266 T; 0 other;  
SQ

Query Match 98.9%; Score 1419.2; DB 22; Length 1459;  
Best Local Similarity 99.7%; Pred. No. 1.8e-254;  
Matches 1432; Conservative 0; Mismatches 3; Indels 1; Gaps 1;

QY 1 CTGGCGGGGCTGGGAACCCAGGCCCCGCGAGGGGCGCCAGAGAGTGGACGCTGGG 60  
|  
Db 1 CTGGCGGGGCTGGGAACCCAGGCCCCGCGAGGGGCGCCAGAGAGTGGACGCTGGG 60  
QY 61 CAAAATGGGCAAGAAAGTGGGTGGGCAAGCATATCTGTTTGTGAGTCTCGCTGGAC 120  
|  
Db 61 CAAAATGGGCAAGAAAGTGGGTGGGCAAGCATATCTGTTTGTGAGTCTCGCTGGAC 120  
QY 121 AAGGTGCTCTGTGGATGCTTACCGGCAACCCAGCAAGAAAGTGGCTGTACAGGGCT 180  
|  
Db 121 AAGGTGCTCTGTGGATGCTTACCGGCAACCCAGCAAGAAAGTGGCTGTACAGGGCT 180  
QY 181 CTGCAAGGCTGCTTGGCAAGAGCGGGGAGCCCGGACGCTGTCAGATGCTGAAGATC 240  
|  
Db 181 CTGCAAGGCTGCTTGGCAAGAGCGGGGAGCCCGGACGCTGTCAGATGCTGAAGATC 240  
QY 241 CACCGCAGCGAACCCGACAGTATCTGCAAGTGCATCTCGGGGCGGACGCTGTGGC 300  
|  
Db 241 CACCGCAGCGAACCCGACAGTATCTGCAAGTGCATCTCGGGGCGGACGCTGTGGC 300  
QY 301 CGCTTCTCTCCGGCTTACCCGAGGGGGGCGTGGCGCGCGCTGTCAGAGAGCTGGCG 360  
|  
Db 301 CGCTTCTCTCCGGCTTACCCGAGGGGGGCGTGGCGCGCGCTGTCAGAGAGCTGGCG 360  
QY 361 GCCGCGCTCGCCACAGCATCGGCTGCGCTGCAACTGG-AGCTGGCGCGCGGGGCCAGCG 419  
|  
Db 361 GCCGCGCTCGCCACAGCATCGGCTGCGCTGCAACTGGATCTGGCGCGCGGGGCCAGCG 420  
QY 420 GCTGAGCGCTTGTGCGGAGCAAGAGAGCGCTGTTTGAATTGCATCTAGCCAGACCC 479  
|  
Db 421 GCTGAGCGCTTGTGCGGAGCAAGAGAGCGCTGTTTGAATTGCATCTAGCCAGACCC 480  
QY 480 CGACCGGCTCGGGATGAAGAACTGGCTGAGCTGGAGGATGCGCTGCGAAATCTGAATG 539  
|  
Db 481 CGACCGGCTCGGGATGAAGAACTGGCTGAGCTGGAGGATGCGCTGCGAAATCTGAATG 540  
QY 540 CGGCTCGGGGGGCGGGGGTGGCGAACGGGAGGTGCGCTTGGGCCCCCTTGGAGCCCCGGT 599  
|  
Db 541 CGGCTCGGGGGGCGGGGGTGGCGAACGGGAGGTGCGCTTGGGCCCCCTTGGAGCCCCGGT 600  
QY 600 GCCCTCTGTGCGAGGTGAAGCCGCGCGCGCGCGCGCACCTGCGCCAGACTTTTCTGT 659  
|  
Db 601 GCCCTCTGTGCGAGGTGAAGCCGCGCGCGCGCGCGCACCTGCGCCAGACTTTTCTGT 660  
QY 660 CCAAGGTCAAGCTGTAGTGAATCGGCGCCTGAGCTTGAAGAACCAACAGAGTTGCGCG 719  
|  
Db 661 CCAAGGTCAAGCTGTAGTGAATCGGCGCCTGAGCTTGAAGAACCAACAGAGTTGCGCG 720  
QY 720 CTGTGGGTCTCAAAATGGCGCAAGTGGGGCGCTCACTGACGAGGAGCTGCGGGGCT 779  
|  
Db 721 CTGTGGGTCTCAAAATGGCGCAAGTGGGGCGCTCACTGACGAGGAGCTGCGGGGCT 780  
QY 780 GCGGAGCCCGGCGCTGGACTGCTGCTGAGTGAAGTGAAGGCGAGGAGTGAAGCA 839  
|  
Db 781 GCGGAGCCCGGCGCTGGACTGCTGCTGAGTGAAGTGAAGGCGAGGAGTGAAGCA 840  
QY 840 GGCTTCAAGCTGCTGCGCGCTTGTGTCAGAGCGGAGGGCGCGCGCTGCAAGCTGCAAG 899  
|  
Db 841 GGCTTCAAGCTGCTGCGCGCTTGTGTCAGAGCGGAGGGCGCGCGCTGCAAGCTGCAAG 900  
QY 900 CCTGTGAGGACATCGAGGAGAACGAGCTCAACAGCTGCGAGAGGACTTGGGGCT 959  
|  
Db 901 CCTGTGAGGACATCGAGGAGAACGAGCTCAACAGCTGCGAGAGGACTTGGGGCT 960  
QY 960 GACCGATCCATGGCGGCGCTGAGTGAAGCAGGGGTGACGAGCTTTTGAAGAACTG 1019  
|  
Db 961 GACCGATCCATGGCGGCGCTGAGTGAAGCAGGGGTGACGAGCTTTTGAAGAACTG 1020  
QY 1020 GATGCGCTTGAAGGTTCTTCTGCGGCTATTGCTGAACCCCTGTCCATCCAGGAACTG 1079  
|  
Db 1021 GATGCGCTTGAAGGTTCTTCTGCGGCTATTGCTGAACCCCTGTCCATCCAGGAACTG 1080



QY 1080 AAACCTCAGCTTGGCTTATCTGCTGAGCCTGCTGAGGAGAGCTGATTCCTTCCCAAGA 1139  
 Db 1081 AAATCTCAGCTTGGCTTATCTGCTGAGCCTGCTGAGGAGAGCTGATTCCTTCCCAAGA 1140  
 QY 1140 GCCGAGCAGCTGGGGGTCATCATTTGGGAGATTCTGCTCAGGTACTTTGATGAGTGTGG 1159  
 Db 1141 GCCGAGCAGCTGGGGGTCATCATTTGGGAGATTCTGCTCAGGTACTTTGATGAGTGTGG 1200  
 QY 1200 GGTGGGGGGGAGCTTGGTGGAGATCAGCCTCAGCTTCTCCATCCCAAGAGCGGGGCTT 1259  
 Db 1201 GGTGGGGGGGAGCTTGGTGGAGATCAGCCTCAGCTTCTCCATCCCAAGAGCGGGGCTT 1260  
 QY 1260 ACAGCCAGCCCTTACAGTTTCACTATGAGACACCTGATCTTGTGTCTGAGACTTCA 1319  
 Db 1261 ACAGCCAGCCCTTACAGTTTCACTATGAGACACCTGATCTTGTGTCTGAGACTTCA 1320  
 QY 1320 TCCCTGGGTGCTGCAAGATATCTCAGTGAAGTAAACAGGATCAATCTTGTCTGCCCCAG 1379  
 Db 1321 TCCCTGGGTGCTGCAAGATATCTCAGTGAAGTAAACAGGATCAATCTTGTCTGCCCCAG 1380  
 QY 1380 CTCACACTCAGCGTGGGAGCCCCGAGATGTTAGCAATGATATAAGTATACAGG 1435  
 Db 1381 CTCACACTCAGCGTGGGAGCCCCGAGATGTTAGCAATGATATAAGTATACAGG 1436  
 RESULT 4  
 AAT41463  
 ID AAT41463 standard; cDNA; 1384 BP.  
 AC AAT41463;  
 AT AAT41463;  
 DT 16-JAN-1997 (first entry)  
 DE Mouse TRADD intracellular signalling protein cDNA.  
 XX Tumour necrosis factor receptor-1 associated death domain protein;  
 KM TRADD; signal transduction; cell growth; cell differentiation;  
 KM apoptosis; gene therapy; ds.  
 OS Mus sp.  
 FH Key Location/Qualifiers  
 FT CDS 1..684  
 FT /\*tag= a  
 PN MO9630404-A1.  
 PD 03-OCT-1996.  
 PF 15-MAR-1996; 96MO-US03591.  
 PR 31-MAR-1995; 95US-0414625.  
 PA (TUL- ) TULARIK INC.  
 PI Goeddel DV, Heu H;  
 XX WPI; 1996-455277/45.  
 DR P-SDB; AAM05530.  
 XX Tumour necrosis factor receptor 1 associated death domain protein  
 PT (TRADD) - used to treat and diagnose diseases associated with  
 PT undesirable cell growth, migration and/or differentiation  
 XX Disclosure; Page 34-35; 42pp; English.  
 PS A cDNA sequence (AAT41463) codes for mouse tumour necrosis factor  
 CC receptor-1 (TNR-R1) associated death domain protein (TRADD) (AAM05530),  
 CC a member of a novel family of intracellular signalling proteins that  
 CC modulate cell growth, differentiation and apoptosis. The cDNA can  
 CC be utilised in recombinant murine TRADD prodn., as a probe to  
 CC detect e.g. the presence of TRADD death domain genes, and in gene  
 CC therapy aplys.

XX SQ Sequence 1384 BP; 308 A; 345 C; 431 G; 300 T; 0 other;  
 Query Match 29.8%; Score 427.6; DB 17; Length 1384;  
 Best Local Similarity 78.5%; Pred. No. 1,6e-70;  
 Matches 540; Conservative 0; Mismatches 139; Indels 9; Gaps 2;  
 QY 302 GCTTCTCCGCGCTTACCGGAGAGGGGGGCTGGCGCGCGGCTGCAAGAGAGCTGGCGG 361  
 Db 2 GCTTCTCCGCGCTTACCGGAGAGGGGGGCTGGCGCGCGGCTGCAAGAGAGCTGGCGG 61  
 QY 362 CCGGCTCGCCCAAGCACTCGGTGCGGCTGCACTGAGTGTGCGCGCGCGCGGCGGCGG 421  
 Db 62 CCGGCTTGGCCCAAGAGAGCGCTGGGTGCAAGTGTGCGGCGGCGGCGGCGGCGGCGG 121  
 QY 422 TGGAGCGTTTGTGCGGAGAGAGAGCGCTGTTGAGTTGCATCTTCAAGCCAGAGCCG 481  
 Db 122 TGGAGCGTTTGTGCGGAGAGAGAGCGCTGTTGAGTTGCATCTTCAAGCCAGAGCCG 181  
 QY 482 ACCGGGCTCGGAGATGAGAACTGGCTGAGCTGAGAGATGGGCTGCGAAATCTGAAGTGG 541  
 Db 182 ACCGGGCTCGGAGAGAGAGAACTGGGAGCTGAGAGATGGGCTGCGAAATCTGAAGTGG 241  
 QY 542 GCTCGGGGGCGCGGGGTGCGGAGCGGAGGTGCTTGGCGCCCTTGCAGAGCCCGGTGC 601  
 Db 242 ACTGCACTGGCCAGGGGTGAGGCACTACAGGTAGCTTGTGAGGTTGCAAGTTCCGGGTTT 301  
 QY 602 CTTCTCTGTGAGAGTGAAGCCCGCGCGCGCCAGCTGCCCCAGACTTTTCTGTTC 661  
 Db 302 CTTCTCCGAGCCGAGAGAA-----ACCACTGCGCGCGCTGCGAGACTTTTCTGTTC 355  
 QY 662 AGGGTCAGCTGTAGTGAATCGGCCCGCTGAGCCTGAAAGGACAAAGACGTTGGCGGCT 721  
 Db 356 AGGGTCAGCTGTAGTGAATCGGCCCGCTGAGCCTGAAAGGACAAAGACGTTGGCGGCT 415  
 QY 722 CTGTGGTCTCAATATGCGAGAGTGGGCGCTCACTGACAGAGGCTGCGGCGCTGC 781  
 Db 416 CGGTGGGTCTCAATATGCGAGAGTGGGCGCTCACTGACAGAGGCTGCGGCGCTGC 475  
 QY 782 GGAACCCGGCGCTGAGTCTGCTGCTGCTTACAGATACAGAGCGGAGCTGTACAGAGG 841  
 Db 476 GAGATCTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 535  
 QY 842 CTTTCAAGCTGCTGCGCGGCTTGTGAGAGCGGAGGCGCGCGCGCGCGCGCGCGCGG 901  
 Db 536 CTTTCAAGCTGCTGCGCGGCTTGTGAGAGCGGAGGCGCGCGCGCGCGCGCGCGCGG 595  
 QY 902 TGGTGAAGCACTCGAGAGAGAGAGTCACTACAGCTGCG---AGAGGACTTGTGCGGCG 958  
 Db 596 TGGTGAAGCGCTGAGAGAGAGAGAGTCACTACAGCTGCG---AGAGGACTTGTGCGGCG 655  
 QY 959 TGACCGATCCCAATGGCGGCTGCGCTTA 986  
 Db 656 AGGCGGAGCCGAGATGGCGGCTGCGCTTA 683  
 RESULT 5  
 ABQ4658  
 ID ABQ4658 standard; DNA; 649 BP.  
 XX ABQ4658;  
 AC ABQ4658;  
 DT 12-JUL-2002 (first entry)  
 XX Oligonucleotide for detecting cytosine methylation SEQ ID NO 33249.  
 DE Human; cytosine methylation; 5'-CpG-3'; uracil; cytosine; diagnosis;  
 KM drug; side effect; cancer; central nervous system; cardiovascular;  
 KM gastrointestinal; respiratory system; single nucleotide polymorphism;  
 KM SNP; cell differentiation; ds.  
 OS Homo sapiens.  
 XX

PN WO200218632-A2.  
 XX  
 PD 07-MAR-2002.  
 XX  
 PF 01-SEP-2001; 2001WO-EP10074.  
 XX  
 PR 01-SEP-2000; 2000DE-1043826.  
 PR 05-SEP-2000; 2000DE-1044543.  
 XX  
 PA (EPIC-) EPIDENOMICS AG.  
 XX  
 PI Olek A, Piepenbrock C, Berlin K, Guetig D;  
 XX WPI; 2002-371829/40.  
 DR  
 XX  
 PT Determining the degree of cytosine methylation in genomic DNA, useful  
 PT for diagnosis and prognosis, comprises selective hybridization of  
 PT amplicons from chemically treated DNA -  
 XX  
 PS Claim 12; 56bp + Sequence Listing; 56bp; German.  
 XX  
 CC This invention describes a novel method for determining the degree of  
 CC methylation of a particular cytosine in a motif 5'-CpG-3', present in a  
 CC genomic sample of DNA. The sample is treated chemically to convert  
 CC cytosine (C) but not methylated C, to uracil, then part of the genomic  
 CC DNA that contains the target C is amplified to form a labeled amplicon.  
 CC The amplicon is hybridised to two classes, each with at least one  
 CC member of oligonucleotides and/or peptide-nucleic acid (PNA) oligomers  
 CC and the degree of hybridisation to both classes is determined from the  
 CC label on the amplicon. From the ratio of labels hybridised to the two  
 CC classes of oligomers, the degree of methylation is calculated. The method  
 CC is used: (i) for diagnosis and/or prognosis of side effects of  
 CC therapeutic drugs and of a wide range of diseases, e.g. cancer, disorders  
 CC of the central nervous, cardiovascular, gastrointestinal and respiratory  
 CC systems etc., particularly by detecting mutations or single nucleotide  
 CC polymorphisms (SNP's); and (ii) for differentiation of cell or tissue  
 CC types and for investigating cell differentiation. The method allows the  
 CC methylation status of many C residues to be determined simultaneously.  
 CC ABQ13410-ABQ54121 represent genomic DNA sequences used to illustrate the  
 CC method for determining the degree of cytosine methylation described in  
 CC the disclosure of the invention.  
 CC  
 XX  
 SQ Sequence 649 BP; 88 A; 74 C; 249 G; 238 T; 0 other;

Query Match 15.2%; Score 238.4; DB 24; Length 649;  
 Best Local Similarity 65.0%; Pred. No. 1e-31;  
 Matches 386; Conservative 0; Mismatches 126; Indels 82; Gaps 1;

QY 165 GGCAGGTGACAGGGCTCTGACAGGCTCTTGACAGAGAGCGGCGAGGCCGCGACGTGCT 224  
 DB 17 GGGATGTTCCGGGTTTGAATCGTTCGTTTCGTAGAGAGCGGCGGAGTTCCGACGTGTT 76  
 QY 225 GCAGATGCTGAAGATCCACCGCAGCGACCGCGACGCTGATCTGCACTGCGATTCGCCG 284  
 DB 77 GTAGAGTTGAAGATTATTCGTAGCCATTCTGATGATCGTGTATTCGATTTTGCCG 136  
 QY 285 GGGGCGACCTGAGCGGCTTCTCCGCGCTACCGCGAGGGGCGCTGCGCGCGCT 344  
 DB 137 GCGGTGATTTTGTGTCGTTTTCGCGTTTATCCGAGGGGCGTTGCGCGCTCGCTT 196  
 QY 345 GCAGAGAGACCTGCGCGCGCGCTCGCCAGCACTCGGTCCGCTCAACTGAGCTGCG 404  
 DB 197 GTAGAGAGAGTTGGCGCGCTCGCTTGTATGATTCGTTGTAATGAGATTGCG 256  
 QY 405 CGCGCGCGCGCGCGCTGACCGCTTTCCTGCGCGAGACGAGAGCCGCTGTTGAGTTGCAT 464  
 DB 257 CGTCGCGCTGACGCGGTTGACGTTTGTGTGGCGAGACGAGACGTTGTTGAGTTGAT 316  
 QY 465 CCTAGCCCAAG----- 474  
 DB 317 TTGAGTTTGAAGTGCCTGCGGTTAGGTTAGATGGGATGAGGCGGAGATCCCGGT 376  
 QY 475 -----CAGCCGACCGCGCTCCGGATGAAGAAC 502

DB 377 TAAGTTTTCGTGACGCGTACGTTTATTTTATTTTGTGATCGGTTTCGGATGAAGAT 436  
 QY 503 TGGCTGAGCTGAGAGATCGCGTCCGCAATCTGAAGTGCAGCTCGGGGCGCGGGTGCAG 562  
 DB 437 TGGTTGAGTTGAGAGATCGCTTGCAGAAATTGAACTGCGGTTCCGGGGGTTCCGGGTGCG 496  
 QY 563 ACGGAGAGTGCCTTCGCGCCCTTGACAGCCCCCGGTGCGCTCTCTGTGAGAGTGAAGC 622  
 DB 497 ACGGAGAGTGCCTTCGCGGTTTGTGAGTTTTCGATGTTTTCGAGAGTGAAGT 556  
 QY 623 CGCGCGCGCGCGCGCGCTGACCGCTGCGAGACTTTCGTTCCAGGTCAGCTTAG 676  
 DB 557 CGTCGCTGCTGCGCTTATTTGATTTTGTGTTAGTTTTCGTTAGGTTAGTTGATG 610

RESULT 6  
 ABQ46659/C  
 ID ABQ46659 standard; DNA; 649 BP.  
 AC  
 AC ABQ46659;  
 XX  
 DT 12-JUL-2002 (first entry)  
 XX  
 DE Oligonucleotide for detecting cytosine methylation SEQ ID NO 33250.  
 XX  
 KW Human; cytosine methylation; 5'-CpG-3'; uracil; cytosine; diagnosis;  
 KW drug; side effect; cancer; central nervous system; cardiovascular;  
 KW gastrointestinal; respiratory system; single nucleotide polymorphism;  
 KW SNP; cell differentiation; de.  
 XX  
 OS Homo sapiens.  
 XX  
 PN WO200218632-A2.  
 XX  
 PD 07-MAR-2002.  
 XX  
 PF 01-SEP-2001; 2001WO-EP10074.  
 XX  
 PR 01-SEP-2000; 2000DE-1043826.  
 PR 05-SEP-2000; 2000DE-1044543.  
 XX  
 PA (EPIC-) EPIDENOMICS AG.  
 XX  
 PI Olek A, Piepenbrock C, Berlin K, Guetig D;  
 XX WPI; 2002-371829/40.  
 DR  
 XX  
 PT Determining the degree of cytosine methylation in genomic DNA, useful  
 PT for diagnosis and prognosis, comprises selective hybridization of  
 PT amplicons from chemically treated DNA -  
 XX  
 PS Claim 12; 56bp + Sequence Listing; 56bp; German.  
 XX  
 CC This invention describes a novel method for determining the degree of  
 CC methylation of a particular cytosine in a motif 5'-CpG-3', present in a  
 CC genomic sample of DNA. The sample is treated chemically to convert  
 CC cytosine (C) but not methylated C, to uracil, then part of the genomic  
 CC DNA that contains the target C is amplified to form a labeled amplicon.  
 CC The amplicon is hybridised to two classes, each with at least one  
 CC member of oligonucleotides and/or peptide-nucleic acid (PNA) oligomers  
 CC and the degree of hybridisation to both classes is determined from the  
 CC label on the amplicon. From the ratio of labels hybridised to the two  
 CC classes of oligomers, the degree of methylation is calculated. The method  
 CC is used: (i) for diagnosis and/or prognosis of side effects of  
 CC therapeutic drugs and of a wide range of diseases, e.g. cancer, disorders  
 CC of the central nervous, cardiovascular, gastrointestinal and respiratory  
 CC systems etc., particularly by detecting mutations or single nucleotide  
 CC polymorphisms (SNP's); and (ii) for differentiation of cell or tissue  
 CC types and for investigating cell differentiation. The method allows the  
 CC methylation status of many C residues to be determined simultaneously.  
 CC ABQ13410-ABQ54121 represent genomic DNA sequences used to illustrate the  
 CC method for determining the degree of cytosine methylation described in

CC the disclosure of the invention.

XX Sequence 649 BP; 238 A; 249 C; 74 G; 88 T; 0 other;

Query Match 15.2%; Score 218.4; DB 24; Length 649;  
Best Local Similarity 65.0%; Pred. No. 1e-31;  
Matches 386; Conservative 0; Mismatches 126; Indels 82; Gaps 1;

```

QY 165 GGCAGTGTACAGGCGCTTCGACGCTGCGCTGGAGAGAGCGGCGGAGCGCGGACGCGT 224
DB 633 GGGATCGTTCGGCGGTTGATCGCTTCGTTTCGTAAGAGAGCGGCGGAGTTGCGACGTTT 574
QY 225 GCAAGTCTGAAGATCCACCGACGACCGCGACGCTGATGCTGACGCTGCGATTTCGCG 284
DB 573 GTAGATGTTGAAATTTATCGTAGCGATTCGATGATGATGATGATGATGATGATGATGAT 514
QY 285 GCGGACGCTGTCGCGCTTCCTCCGCGCTACCGGAGAGGCGCGCTGCGCGCGCT 344
DB 513 GCGGTAGTTCGTCGCTTCGCTTCGCTTCGCTTCGCTTCGCTTCGCTTCGCTTCGCTTC 454
QY 345 GCAAGAGAGCGCTGCGCGCGCTGCGCGCTGCGCGACACTGCGCTGCGACGCTGAGCTGCG 404
DB 453 GTAGAGAGATTTGCGCGCTTCGCTTCGCTTCGCTTCGCTTCGCTTCGCTTCGCTTCG 394
QY 405 GCGCGCGCGCGAGCGCTGCAAGCTTCGCTGCGCGAGAGAGCGCTGTTGAGTTGCAT 464
DB 393 GGTGCGGTGCGAGCGGTTGACGTTTGTGCGCGAGAGAGCGGTTGTTGAGTTGAT 334
QY 465 CTTAGCCGAC----- 474
DB 333 TTTAGTTTAAAGTGGCGGTGCGGTGAGGTTAGGTTAGGTTAGGTTAGGTTAGGTTAGG 274
QY 475 -----CAGCCGACCGGCTCCGCGGATGAAGAC 502
DB 273 TAACTTTTGTACGCTAGCTTTATTTTATTTTATTTTATTTTATTTTATTTTATTTTAT 214
QY 503 TGCGTGAAGTGAAGATGCGCTGCGAAATCTGAAGTGCAGCTGCGGCGCGCGGCTGCGG 562
DB 213 TGGTTGATTTGAGAGATGCGTTGCGAAATTTGAAGTGCAGGTTGCGGCGGTTGCGG 154
QY 563 ACGGAGAGTTCGCTTCGCGCGCGCTTCGAGCGCGCGCGCTTCCTGCTGCGAGGTGAAGC 622
DB 153 ACGGAGAGTTCGCTTCGCTTCGCTTCGCTTCGCTTCGCTTCGCTTCGCTTCGCTTCG 94
QY 623 GCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 676
DB 93 GGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 40

```

RESULT 7

ABQ4660/C

ID ABQ4660 standard; DNA; 649 BP.

XX ABQ4660:

DT 12-JUL-2002 (first entry)

DE Oligonucleotide for detecting cytosine methylation SEQ ID NO 33251.

XX Human; cytosine methylation; 5'-CpG-3'; uracil; cytosine; diagnosis;  
KW drug; side effect; cancer; central nervous system; cardiovascular;  
KM gastrointestinal; respiratory system; single nucleotide polymorphism;  
SNP; cell differentiation; db.

OS Homo sapiens.

XX MO200218632-A2.

PD 07-MAR-2002.

PF 01-SEP-2001; 2001MO-EP10074.

XX 01-SEP-2000; 2000DE-1043826.

PR 05-SEP-2000; 2000DE-1044543.

XX (EPIC-) EPIGENOMICS AG.

PI Olek A, Piepenbrock C, Berlin K, Guetig D;

DR WPI; 2002-371829/40.

PT Determining the degree of cytosine methylation in genomic DNA, useful  
for diagnosis and prognosis, comprises selective hybridization of  
PT amplicons from chemically treated DNA

PS Claim 12; 56pp + Sequence Listing; 56pp; German.

CC This invention describes a novel method for determining the degree of  
CC methylation of a particular cytosine in a motif 5'-CpG-3', present in a  
CC genomic sample of DNA. The sample is treated chemically to convert  
CC cytosine (C) but not methylated C, to uracil, then part of the genomic  
CC DNA that contains the target C is amplified to form a labeled amplicon.  
CC The amplicon is hybridized to two classes, each with at least one  
CC member, of oligonucleotides and/or peptide-nucleic acid (PNA) oligomers  
CC and the degree of hybridization to both classes is determined from the  
CC label on the amplicon. From the ratio of labels hybridized to the two  
CC classes of oligomers, the degree of methylation is calculated. The method  
CC is used: (i) for diagnosis and/or prognosis of side effects of  
CC therapeutic drugs and of a wide range of diseases, e.g. cancer, disorders  
CC of the central nervous, cardiovascular, gastrointestinal and respiratory  
CC systems etc., particularly by detecting mutations or single nucleotide  
CC polymorphisms (SNP's) and (ii) for differentiation of cell or tissue  
CC types and for investigating cell differentiation. The method allows the  
CC methylation status of many C residues to be determined simultaneously.  
CC ABQ13410-ABQ54121 represent genomic DNA sequences used to illustrate the  
CC method for determining the degree of cytosine methylation described in  
CC the disclosure of the invention.

XX Sequence 649 BP; 98 A; 74 C; 214 G; 263 T; 0 other;

Query Match 13.9%; Score 198.8; DB 24; Length 649;  
Best Local Similarity 63.4%; Pred. No. 4.4e-28;  
Matches 370; Conservative 0; Mismatches 132; Indels 82; Gaps 1;

```

QY 184 CAGGCTGCTTGGCAGAGAGCGGCGGAGCGCGAGCTGCTGAGATGCTGAATTCAC 243
DB 614 CCGCCGCGCTCCGCAAAAACGCAAAAACCGACGTAACAAATCTAAATAATCCAC 555
QY 244 CGAGAGAGAGCGGAGCTGATGCTGAGCTGCGATTTCGCGGCGCGAGCCCTTGCGCGC 303
DB 554 CGCAAGAGAGCGGCACTAATCGTCAAACTAAGATTCTACGAAACGACAACTTAACCGC 495
QY 304 TTCCTCGCGGCTTACCGCGAGGAGGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 363
DB 494 TTCCTCGCGGCTTACCGCGAGAAAGCTTACGCGCGCGCTTACAAATAATCTTAAGACC 435
QY 364 GCGCTCGCGGAGCACTCGTGCCTGCTGCAACTGAGCTGCGCGCGCGCGCGCGCGCTG 423
DB 434 GCGCTCGCGGAGCACTCGTGCCTGCTGCAACTGAGCTGCGCGCGCGCGCGCGCGCGCT 375
QY 424 GAGGCTTTCGCTGCGGAGAGAGCGCTGTTGATGATCTTAAGCCAGC----- 475
DB 374 AAGGCTTTCGCTGCGGAGAGAGAGCGCTGTTGATGATCTTAAGCCAGCAATAAGCA 315
QY 476 ----- 475
DB 314 CCGAATTAATAATATAATAATAATAATAATAATAATAATAATAATAATAATAATAATA 255
QY 476 -----AGCCGAGCGGCTCGGAGTGAAGAATGCTGAGCTGAGAGATGC 521
DB 254 GGTTCATTCCTTAACCGAGCGGAGTGAAGAATTAATAATAATAATAATAATAATAATAC 195
QY 522 GCTGCGAAATCTGAAGTGGCGGCTCGGCGGCTCGGCGGCTCGGCGGAGGTGCTTCGCG 561
DB 194 GCTACGAATCTTAATAATGACATCGAAGAACCCGAAATAACGAGGAAATAATCGCTTCGAC 135

```

OY 582 CCCCTTGAGCCCCCGGTGCTCTGTGTGAGAGTGAAGCCCGCCGCGCCGAC 641  
 DB 134 CCCCTTACACCCCGGATACCTCTATCGAAAAATAAACCGCGCCGCGCAC 75  
 OY 642 TGCCGAGCTTTTCTGTCGAGGTGAGCTGTAGTAATGCGC 685  
 DB 74 TACCCAACTTTCTATTCAAAATCACTAATTAATAACGC 31

RESULT 8  
 ABQ46661  
 ID ABQ46661 standard; DNA; 649 BP.

ABQ46661;  
 12-JUL-2002 (first entry)

Oligonucleotide for detecting cytosine methylation SEQ ID NO 33252.

Human; cytosine methylation; 5'-CpG-3'; uracil; cytosine; diagnosis;  
 drug; side effect; cancer; central nervous system; cardiovascular;  
 gastrointestinal; respiratory system; single nucleotide polymorphism;  
 SNP; cell differentiation; ds.

Homo sapiens.

WO200218632-A2.

07-MAR-2002.

01-SEP-2001; 2001WO-EP10074.

01-SEP-2000; 2000DE-1043826.

05-SEP-2000; 2000DE-1044543.

(EPIC-) EPIDENOMICS AG.

Olek A, Piepenbrock C, Berlin K, Guetig D;  
 WPI, 2002-371829/40.

Determining the degree of cytosine methylation in genomic DNA, useful  
 for diagnosis and prognosis, comprises selective hybridization of  
 amplicons from chemically treated DNA -

Claim 12; 56pp + Sequence Listing; 56pp; German.

This invention describes a novel method for determining the degree of  
 methylation of a particular cytosine in a motif 5'-CpG-3', present in a  
 genomic sample of DNA. The sample is treated chemically to convert  
 cytosine (C) but not methylated C, to uracil, then part of the genomic  
 DNA that contains the target C is amplified to form a labeled amplicon.  
 The amplicon is hybridised to two classes, each with at least one  
 member, of oligonucleotides and/or peptide-nucleic acid (PNA) oligomers  
 and the degree of hybridisation to both classes is determined from the  
 label on the amplicon. From the ratio of labels hybridised to the two  
 classes of oligomers, the degree of methylation is calculated. The method  
 is used: (i) for diagnosis and/or prognosis of side effects of  
 therapeutic drugs and of a wide range of diseases, e.g. cancer, disorders  
 of the central nervous, cardiovascular, gastrointestinal and respiratory  
 systems etc., particularly by detecting mutations or single nucleotide  
 polymorphisms (SNP's); and (ii) for differentiation of cell or tissue  
 types and for investigating cell differentiation. The method allows the  
 methylation status of many C residues to be determined simultaneously.  
 ABOJ3410-ABQ5121 represent genomic DNA sequences used to illustrate the  
 method for determining the degree of cytosine methylation described in  
 the disclosure of the invention.

Sequence 649 BP; 263 A; 214 C; 74 G; 98 T; 0 other;

Query Match 13.9%; Score 198.8; DB 24; Length 649;  
 Best Local Similarity 63.4%; Pred. No. 4.4e-28;  
 Matches 370; Conservative 0; Mismatches 132; Indels 82; Gaps 1;

OY 184 CAGGCTGCTTTGGCAGAGAGGCGGGAGCCCGGACGTGCTGAGATGCTGAATCCAC 243  
 DB 36 CCGCGCGCTCCGCAAAAAAGAGAAAAACCGAAGCTATCAAAATACTAAAAATCCAC 95  
 OY 244 CGCAGCAGCCCGGAGCTGATGTGAGGTGCTTTGCGGGCGGCGACCTGTGCGCG 303  
 DB 96 CGCAGCAGCCCGGAGCTGATGTGAGGTGCTTTGCGGGCGGCGACCTGTGCGCG 155  
 OY 304 TTCTCCGCGCTTACCGGAGAGGGGCGTGGCGCGCGGTGACAGAGAGCTGAGCGCC 363  
 DB 156 TTCTCCGCGCTTACCGGAGAGGGGCGTGGCGCGCGGTGACAGAGAGCTGAGCGCC 215  
 OY 364 GCGCTGCGCCAGCACTCGGTGCGGTGCGAATGAGAGCTGCGCGCGCGCGCGCGT 423  
 DB 216 GCGCTGCGCCAGCACTCGGTGCGGTGCGAATGAGAGCTGCGCGCGCGCGCGCGT 275  
 OY 424 GACGCTTGTGGGCGAGAGAGGCGGTGAGTGTGACCTAGCCAGC----- 475  
 DB 276 AACGCTTACTAG 335  
 OY 476 ----- 475  
 DB 336 CCGAAGTAAATCAAAATATAAATAAAGAGAAATCCGAGTTAAATCCCGCTAACACA 395  
 OY 476 -----AGCCGACCGGCTCGGATGAGAACTGGCTGAGATGC 521  
 DB 396 GGTTCATTCCTTACCCGATACCGGACCGATCCGAAATATAAATACTAAATCAAAATATAC 455  
 OY 522 GGTGGAATCTGAAGTGGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 581  
 DB 456 GCTACGAATCTGAAGTGGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 515  
 OY 582 CCCCTTGAGCCCCCGGTGCTCTGTGTGAGAGTGAAGCCCGCCGCGCCGAC 641  
 DB 516 CCCCTTACACCCCGGATACCTCTATCGAAAAATAAACCGCGCGCGCGCAC 575  
 OY 642 TGCCGAGCTTTTCTGTCGAGGTGAGCTGTAGTAATGCGC 685  
 DB 576 TACCCAACTTTCTATTCAAAATCACTAATTAATAACGC 619

RESULT 9  
 AAQ98470  
 ID AAQ98470 standard; cDNA; 2744 BP.

AAQ98470;

01-MAY-1996 (first entry)

MISP1-containing plasmid pMIS1.

Spider silk; repeat unit; consensus; minor ampullate silk protein;  
 spider; orb web spider; dragline; ds.

Nephila clavipes.

Key Location/Qualifiers

FT CDS 183..2678

FT Misc-difference 996..1037

FT FT /\*tag= b

FT FT /\*note= "represented as indeterminate, a highly

FT FT compressed GC rich region which could

FT FT not be sequenced"

PN WO9525165-A1.

PD 21-SEP-1995.

PF 14-MAR-1995; 95WO-US03139.

PR 14-MAR-1994; 94US-0209747.

XX (UYWY-) UNIV WYOMING.  
 PA Colgin M, Lewis RV;  
 PI WPI, 1995-336970/43.  
 DR P-PSDB; AAR80168.  
 XX Polypeptide(s) comprising repeated unit amino acid sequences, also  
 PT cDNAs - derived from minor ampullate spider silk proteins and used  
 PT to form spider silk fibres  
 XX Claim 12, Fig 1; 86pp; English.  
 XX cDNA clone, pMIS1, encoding the orb web spider minor ampullate silk  
 CC protein Misp1 has been identified and sequenced. Repeat unit peptides  
 CC (see AAR80184-85) and a carboxy terminus (see AAR80186-89). Repeat unit  
 CC peptides make up spider silk proteins (spidroins) which in turn  
 CC aggregate to form the silk fibres. Spider silk fibres have high  
 CC tensile strength and significant elasticity. An isolated cDNA clone of  
 CC a silk protein encoding sequence is of use to produce the protein at  
 CC high yields using recombinant DNA technology.  
 XX Sequence 2744 BP; 486 A; 549 C; 1059 G; 608 T; 42 other;  
 SQ  
 Query Match 5.0%; Score 71.2; DB 16; Length 2744;  
 Best Local Similarity 43.8%; Pred. No. 0.00022;  
 Matches 362; Conservative 0; Mismatches 458; Indels 6; Gaps 1;  
 QY 104 TGGAGTCTCGCTGAGCAAGGTGTCTCTGTCGATGCTTACGCCACCCCGAGCAAGG 163  
 DB 1199 TGGAGGCTAGCGTGTGTAAGTGTATACGTCGCCGAGCAGAGACTGTGAGCTGCTGG 1258  
 QY 164 TGGCAGTGTACAGGCGCTCTGCAAGCTGCTTGGCAGAGAGCGCGGAGACCTGC 223  
 DB 1259 AGCAGAGCTGAGAGCGCTGTGTGTACGTCGAGTGTCTGTCTGAGCAGAGCGCG 1318  
 QY 224 TGCAGATGCTGAAGATCCACCGCAGCGACCGGAGCTGATGTGACCTGTGATTTGGC 283  
 DB 1319 TGCAGTGTCTGAGCAGAGAGCGCTGCGGAGTCAGAGACTGAGAGCTACGGTGTCAAG 1378  
 QY 284 GCGCGCAGCCCTGTGCGCGCTTCTCCGCGCTTACCGCGAGGGGCGCTGCGCGCGC 343  
 DB 1379 TGGGTAGCGTCCGCTGAGAGAGCTGTGCGCGCTGTCTGTCTGAGCAGAGAGCTGAG 1438  
 QY 344 TGCAGAGAGCTGCGCGCGCTGCGCGCAGCACTGCGTCCGCTGCACTGAGCTGC 403  
 DB 1439 CGCTGTGTGTACCGTGTAGAGGTGCTGTGTGTGAGCTGAGCTGTCAAGCGCAGAGC 1498  
 QY 404 GCGCGCGCGCGCAGCGCTGAGACCTTGTCTGCGGAGCAGAGAGCGCTGTGTGAGTTCA 463  
 DB 1499 TGGAGGCTACCGTGTCAAGGTGATACGTCGCGAGCAGAGACTGTGTGTGAGC 1558  
 QY 464 TCCTAGCCCAAGCAGCCGAGCGCTCCGCGAT-----GAAGAACTGGCTGAGCTGAG 517  
 DB 1559 TGCCTGCAACAGAGAGCGCGAGCGCTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 1618  
 QY 518 ATGCGCTTGCAGAAATCTGAAGTGTGCGCTGTGCGGCGCGCGGAGTGTGCTT 577  
 DB 1619 TGCCTGT 1678  
 QY 578 GCGCGCGCGCGCGCGCGCGCTGTCTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 637  
 DB 1679 TGGT 1738  
 QY 638 CACCTGCCCAAGCTTTTCTGTTCAGAGGCTACGCTGTGTGTGTGTGTGTGTGTGTGTGT 697  
 DB 1739 TTACGTAAGAGT 1798  
 QY 698 AGGACCAAGAGCTTGT 757  
 DB 1799 TGCAGGT 1858

QY 758 TGCAGCAGGCTCGCGGCGCTGCGGAGCCCGGCGTGTGACTGTGCTTACGAGTACG 817  
 DB 1859 AAGAGCTGT 1918  
 QY 818 AGCGCGAGGAGCTGTACGAGCAGCGCTTCCAGCTGTGTGCGCGCTTGTGTGAGCCGAG 877  
 DB 1919 AGGT 1978  
 QY 878 GCGCGCGCGCGCGCGCTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 923  
 DB 1979 AGCTGT 2024  
 RESULT 10  
 AAX53491/C  
 ID AAX53491 standard; DNA; 114955 BP.  
 XX  
 AC AAX53491;  
 XX  
 DT 05-JUL-1999 (first entry)  
 XX  
 DE Human adenosine A1 receptor antisense oligonucleotide fragment.  
 XX  
 KW Antisense oligonucleotide; multiple target; antisense treatment;  
 KW impaired respiration; inflammation; lung disease;  
 KW pulmonary vasoconstriction; inflammation; allergic rhinitis;  
 KW acute asthma; allergy; asthma; impeded respiration;  
 KW respiratory distress syndrome; pain; cystic fibrosis;  
 KW pulmonary hypertension; pulmonary vasoconstriction; emphysema;  
 KW chronic obstructive pulmonary disease; leukemia; lymphoma; carcinoma;  
 KW colon cancer; breast cancer; lung cancer; pancreatic cancer;  
 KW hepatocellular carcinoma; kidney cancer; melanoma; hepatic metastasis;  
 KW prostate cancer; ss.  
 XX  
 OS Synthetic.  
 XX  
 PN WO9913886-A1.  
 XX  
 PD 25-MAR-1999.  
 XX  
 PF 17-SEP-1998; 98WO-US19419.  
 XX  
 PR 09-JUN-1998; 98US-0093972.  
 PR 17-SEP-1997; 97US-0059160.  
 XX  
 PA (UYEC-) UNIV EAST CAROLINA.  
 XX  
 PI Nyce JW;  
 XX  
 DR WPI, 1999-229400/19.  
 XX  
 PT New antisense oligonucleotides used in treatment of, e.g. pulmonary  
 PT vasoconstriction  
 PS  
 XX Disclosure; Page 37; 120pp; English.  
 XX  
 CC The specification describes antisense oligonucleotides (AAX52869-X55271)  
 CC directed against at least 2 mRNAs selected from target genes, coding and  
 CC non-coding regions of RNAs corresponding to target genes, gene  
 CC initiation codons, genomic flanking regions, intron-exon borders, the  
 CC 5'-end, the 3'-end and the juxta-section between coding and non-coding  
 CC regions and all segments of RNAs encoding proteins associated with one  
 CC or more diseases, conditions or mixtures. The antisense oligonucleotides  
 CC may be derived from sequences AAX55180-271. These multiple target  
 CC oligonucleotides (specifically AAX55180-271) can be used for the  
 CC antisense treatment of diseases and conditions. Typical diseases and  
 CC conditions are those associated with impaired respiration and  
 CC inflammation, including lung diseases, pulmonary vasoconstriction,  
 CC inflammation, allergic rhinitis, acute asthma, allergies, asthma, impeded  
 CC respiration, respiratory distress syndrome, pain, cystic fibrosis,  
 CC pulmonary hypertension, pulmonary vasoconstriction, emphysema, chronic  
 CC obstructive pulmonary disease (COPD), and cancers such as leukemias,

CC lymphomas, carcinomas e.g. colon cancer, breast cancer, lung cancer,  
 CC pancreatic cancer, hepatocellular carcinoma, kidney cancer, melanoma,  
 CC hepatic metastases, as well as all types of cancers which may metastasize  
 CC or have metastasized to the lungs, including breast and prostate cancer.

XX Sequence 114955 BP; 6071 A; 29417 C; 36712 G; 21328 T; 21427 other;

Query Match 4.8%; Score 69.2; DB 20; Length 114955;  
 Best Local Similarity 31.9%; Pred. No. 0.0006;  
 Matches 266; Conservative 83; Mismatches 477; Indels 9; Gaps 2;

QY 165 GGCAGTGTACAGAGCTCTGACAGCTGCTTTGGCAGAGAGCGCGGAGCGCCGACGTCT 224

DB 104951 GCGCGCGCGCCCG 104892

QY 225 GCAGATGTCTGAAGATCCACCGCAGCGACCTGATCTGTGCACTCCGATTCTGCGCG 284

DB 104891 HNNNSGGCCGCGGNNNNNSVGGCCGCGNNNNNSCVGGCGGNNNNNSCVGG 104832

QY 285 GCGGAGAGCCCTGTGGCGGCTTCTCTCCGCGCTACCGCGAGAGGAGCGCTGCGCGCGCT 344

DB 104831 GCGCGCGGNNNNNSGCCGCGCGCGCGGNNNNNSGCCGCGCGCGGNNNNNSGCC 104772

QY 345 GCAGAGAGAGCTGTGGCGCGCGCGCTGCGCGACGCTGCTGCGCGCTGCGCGCTGCG 404

DB 104771 CVGGCGCGGNNNNNSGCCGCGCGCGCGGNNNNNSGCCGCGCGCGCGCGGNN 104712

QY 405 CGCGCGCGCGCGCGCGCTGTGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 464

DB 104711 HNNNSGGCG 104652

QY 465 CCTAGCCCGACAGCCCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCT 524

DB 104651 GCG 104592

QY 525 GCGAATCTGAAGTGTGCGCGCTGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCT 584

DB 104591 CCG 104532

QY 585 CTTCGAGCGCGCGCGCGCTCTCTCTGTGCGAGTGAAGCGCGCGCGCGCGCGCGCT 644

DB 104531 GCG 104472

QY 645 CCAGACTTTTCTGTTCCAGAGGTGCTAGTGAATCGGCGCGCTGAGCTGAAGGACCA 704

DB 104471 GCG 104412

QY 705 ACAGAGCTTGTGCGCGCTCTGTGAGTCAATGAGCGCAAGTTGG-GGCGCTCACTGACG 763

DB 104411 CCG 104352

QY 764 GAGGCTGCGCGCGCGCTGCGCGGACCCGCGCGCTGACCTGCGTGGCTCAAGATAC----- 816

DB 104351 GCGGNNNNNSGCCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 104292

QY 817 -GAGCGCGAGGAGCTGTACGAGGAGCTTCCAGCTGTGCGCGCGCTTGTGAGGCGCA 875

DB 104291 GCG 104232

QY 876 GAGCGCGCGCGCGCGCGCTGCGAGCGCGCTGTGAGGAGCTGAGAGAGAGAGCTACCA 935

DB 104231 HNNNSGCCGCGCGCGGNNNNNSGCCGCGCGCGGNNNNNSGCCGCGCGGCA 104172

QY 936 CCGCGCGAGAGAGCTTGTGAGGCTGACCGAGTCCCAATGGCGCGCGCTGAGCTTAC 990

DB 104171 CCGCGCGCGCGCGGNNNNNSGCCGCGCGCGCGCGCGCGCGCGCGCGCGCG 104117

RESULT 11  
 AAX53491  
 ID AAX53491 standard; DNA; 114955 BP.  
 XX  
 AC AAX53491;

XX 05-JUL-1999 (first entry)  
 DT  
 XX  
 DE Human adenosine A1 receptor antisense oligonucleotide fragment.

KW Antisense oligonucleotide; multiple target; antisense treatment;  
 KW impaired respiration; inflammation; lung disease;  
 KW pulmonary vasoconstriction; inflammation; allergic rhinitis;  
 KW acute asthma; allergy; asthma; impeded respiration;  
 KW respiratory distress syndrome; pain; cystic fibrosis;  
 KW pulmonary hypertension; pulmonary vasoconstriction; emphysema;  
 KW chronic obstructive pulmonary disease; leukemia; lymphoma; carcinoma;  
 KW colon cancer; breast cancer; lung cancer; pancreatic cancer;  
 KW hepatocellular carcinoma; kidney cancer; melanoma; hepatic metastasis;  
 KW prostate cancer; ss.

OS Synthetic.

PN WO913886-A1.

PD 25-MAR-1999.

PF 17-SEP-1998; 98WO-US19419.

PR 09-JUN-1998; 98US-0093972.

PR 17-SEP-1997; 97US-0059160.

PA (UVEC-) UNIV EAST CAROLINA.

PI Nyce JW;

DR WPI; 1999-229400/19.

PT New antisense oligonucleotides used in treatment of, e.g. pulmonary

PT vasoconstriction

PS Disclosure; Page 37; 120pp; English.

CC The specification describes antisense oligonucleotides (AAX52869-X55271)  
 CC directed against at least 2 mRNAs selected from target genes, coding and  
 CC non-coding regions of RNAs corresponding to target genes, gene  
 CC initiation codons, genomic flanking regions, intron-exon borders, the  
 CC 5'-end, the 3'-end and the juxta-section between coding and non-coding  
 CC regions and all segments of RNAs encoding proteins associated with one  
 CC or more diseases, conditions or mixtures. The antisense oligonucleotides  
 CC may be derived from sequences AAX55272-74. These multiple target  
 CC oligonucleotides (specifically AAX55180-271) can be used for the  
 CC antisense treatment of diseases and conditions. Typical diseases and  
 CC conditions are those associated with impaired respiration and  
 CC inflammation, including lung diseases, pulmonary vasoconstriction,  
 CC inflammation, allergic rhinitis, acute asthma, allergies, asthma, impeded  
 CC respiration, respiratory distress syndrome, pain, cystic fibrosis,  
 CC pulmonary hypertension, pulmonary vasoconstriction, emphysema, chronic  
 CC obstructive pulmonary disease (COPD), and cancers such as leukemias,  
 CC lymphomas, carcinomas e.g. colon cancer, breast cancer, lung cancer,  
 CC pancreatic cancer, hepatocellular carcinoma, kidney cancer, melanoma,  
 CC hepatic metastases, as well as all types of cancers which may metastasize  
 CC or have metastasized to the lungs, including breast and prostate cancer.

XX Sequence 114955 BP; 6071 A; 29417 C; 36712 G; 21328 T; 21427 other;

Query Match 4.5%; Score 64.4; DB 20; Length 114955;  
 Best Local Similarity 32.0%; Pred. No. 0.0047;  
 Matches 318; Conservative 91; Mismatches 576; Indels 10; Gaps 4;

QY 18 CAGAGCCCGCGGAGCGCGCAGAGGTGAGTGCAGCTGGCGCAAAATGGGACAGAA 77

DB 104556 CCBGGGCGCGCGCGGNNNNNSCCBGGCGCGCGCGCGCGGNNNNNSCCBGGC 104615

QY 78 GTGGGTGGCGAGGATATCTGTTGTGAGTCTGCTGAGCAAGTGTCTGTGCGA 137

DB 104616 CBGGGCGCGCGGNNNNNSCCBGGCGCGCGCGCGGNNNNNSCCBGGCGCGG 104675

QY 138 TGCCTACGCGCACCCCGCAGAGAGTGCGATGAGGCTCTGAGAGCTTGCGCTTGCC 197  
 Db 104676 CGCGCCSNNNDNNCCGCGCGCCGCGCCSNNNDNNCCGCGCGCGCGCGSNNND 104735  
 QY 198 AGAGAGGGGGGAGCGCGAGCGTGTGCGAGATGCTGAAGATCCAGCCGACCGCCGCA 257  
 Db 104736 NNC CG 104795  
 QY 258 GCTGATGCTGAGCTGCGATTTCTGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCTA 317  
 Db 104796 CSNNNDNNCCGS 104855  
 QY 318 CGCGAGAGGGGCGCGTGTGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCA 377  
 Db 104856 NNDNNCCG 104915  
 QY 378 CTGCGTCCGCGCTGCACTGAGCTGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCTT 437  
 Db 104916 GCG 104975  
 QY 438 GAGAGAGAGCGCGCTGTGATGCTGATCCTAGCCGAGCGCGCGCGCGCGCGCGCGG 494  
 Db 104976 GGGCG 105035  
 QY 495 TGAAGAACTGCTGAGCTGAGAGATGCGCTGCGAAT---CTGAAGTGTGCGCTGCG 550  
 Db 105036 NDNNGGCTBGGGGGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 105095  
 QY 551 CCGGGGCTGCGCGAGCGGAGAGTCTTGCGCGCGCGCGCGCGCGCGCGCGCGCGCTT 610  
 Db 105096 GCGGCG 105155  
 QY 611 CGGAGGTGAAGCG 670  
 Db 105156 CGGCG 105215  
 QY 671 CTGTAGTGAATGCG 730  
 Db 105216 GGGCG 105275  
 QY 731 TCAATGCGCGCGAGAGTGTGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 788  
 Db 105276 GSNNDNNCCG 105335  
 QY 789 GGGCGGTGAATGCG 848  
 Db 105336 GCG 105395  
 QY 849 GCTGCTGCGCGCGCTTCTGTGAGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 907  
 Db 105396 GSNNDNNCCG 105455  
 QY 908 AGGCACTGAT 967  
 Db 105456 GSNNDNNCCG 105515  
 QY 968 CCAATGCGCGCGCTGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1002  
 Db 105516 NNDNNCCG 105550

RESULT 12  
 AAV55831  
 ID AAV55831 standard; DNA; 799 BP.  
 XX AAV55831;  
 AC  
 XX  
 DT 18-NOV-1998 (first entry)  
 XX  
 DE Nucleotide sequence of the stabilising sequence-encoding insert.  
 XX Fusion protein; stabilising polypeptide; proteolytic degradation;  
 KW resistance; half-life; autoimmune disease; inflammation; nitro drug;

KW Ikappab regulator protein; inflammatory bowel disease; in vivo imaging;  
 KW nitroreductase protein; enzyme therapy; prodrug therapy; protease;  
 KW cancer; pathological condition; ss.  
 XX  
 OS Epstein-Barr virus.  
 XX  
 PN W0982577-A1.  
 XX  
 PD 28-MAY-1998.  
 XX  
 PF 17-NOV-1997; 97WO-IB01508.  
 XX  
 PR 25-JUN-1997; 97US-0048945.  
 PR 15-NOV-1996; 96US-0030986.  
 XX  
 PA (MAST/) MASTUCCI M G.  
 XX  
 PI Masucci MG;  
 XX  
 DR WPI, 1998-312463/27.  
 XX  
 PT New fusion proteins resistant to proteolytic degradation -  
 PT comprising a core protein with a stabilising polypeptide comprising  
 PT a peptide sequence containing glycine repeats  
 XX  
 PS  
 XX  
 CC This is a nucleotide sequence of the stabilising sequence-encoding  
 CC insert. The invention provides a method for increasing the resistance  
 CC of a core protein to proteolytic degradation that comprises linking or  
 CC inserting onto or into the core protein a stabilising polypeptide of  
 CC formula (Gly)(X)(Gly)(Z)n where Gly, Glyb, Glyc are 1-6  
 CC sequential Gly residues and X, Y, Z are Ala, Ser, Val, Ile, Leu, Met,  
 CC Phe, Pro or Thr and n can be anything between 1-66. X, Y and Z need not  
 CC be identical from n repeat to n repeat. Alternatively a nucleic acid  
 CC encoding the stabilising polypeptide can be linked onto or inserted into  
 CC a nucleic acid encoding a core protein. The fusion proteins of the  
 CC invention are more resistant to degradation by proteases and, thus, have  
 CC a longer half-life than the unfused core protein. The products can be  
 CC used for treating autoimmune diseases, cancer and inflammation. In  
 CC particular, the core protein may be an Ikappab regulator protein for the  
 CC treatment of inflammatory bowel disease, or a nitroreductase protein  
 CC which can activate nitro drugs in enzyme/prodrug therapy to treat cancer  
 CC or other pathological conditions. The fusion proteins can also be used in  
 CC diagnostic methods such as in vivo imaging.  
 XX  
 SQ Sequence 799 BP; 201 A; 106 C; 479 G; 13 T; 0 other;

Query Match 4.4%; Score 63.8; DB 19; Length 799;  
 Best local similarity 42.8%; Pred. No. 0.0049;  
 Matches 320; Conservative 0; Mismatches 427; Indels 0; Gaps 0;

QY 103 GTGAGTCTCTGCTGAGCAAGTGTCTGTGATGCTTACGCGAGCCCGCAGAGAG 162  
 Db 18 GTGAGCAAGAGCAG 77  
 QY 163 GTGCACTGTACAGAGGCTCTGTGAGGCTGTGTGAGAGAGCGCGGAGCCCGAGAG 222  
 Db 78 GAGGAGCAG 137  
 QY 223 CTGAGATGCTGAAGATCCAGCGCAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAT 282  
 Db 138 GAGGAGGAG 197  
 QY 283 GGGCGAGCGCTGTGAGCGCGCTTCTCGCGCGCTTACCGGAGAGAGAGAGAGAGAG 342  
 Db 198 GGGCAG 257  
 QY 343 CTGCAAG 402  
 Db 258 GAGGAGCAG 317  
 QY 403 CGGCGCGCGCGAG 462



```

Db 318 GAGGAGGGGACAGAGGGGACAGAGGGGACAGAGGGGACAGAGGGGACAGAGGGG 377
Qy 463 ATCTTACCCACACCCACCCGCTCCCGGATGAAAGACTGGAGTGAAGATGCG 522
Db 378 CAGAGGGGACAGAGGGGACAGAGGGGACAGAGGGGACAGAGGGGACAGAGGGG 437
Qy 523 CTGCAAAATCTGAATGCGGCTCGGGGGCCCGGGGTGCGAGCGGGAGGTCCCTTCGGCC 582
Db 438 CAGAGGGGACAGAGGGGACAGAGGGGACAGAGGGGACAGAGGGGACAGAGGGG 497
Qy 563 CCCTTGACGCCCCCGGTGCTCTCTGTGTGAGGTGAGCCGCCGCCGCCCACT 642
Db 498 CAGAGGGGACAGAGGGGACAGAGGGGACAGAGGGGACAGAGGGGACAGAGGGG 557
Qy 643 GCCCAGACTTTCTGTTCAGGGGTCAAGCTGTAAGTAATCGGCCCTGAGCCTGAAGAC 702
Db 558 GAGGGGACAGAGGGGACAGAGGGGACAGAGGGGACAGAGGGGACAGAGGGG 617
Qy 703 CAACAGACGTTGCGCGCTCTGTGGGTCTCAATGCGCAGAGGTGGGCGCTCACTGCAG 762
Db 618 CAGAGGGGACAGAGGGGACAGAGGGGACAGAGGGGACAGAGGGGACAGAGGGG 677
Qy 763 CGAGGCTCGCGGCGCTGTGCGGAAACCCGCGCTGGAATCGCTGAGTACGAGCC 822
Db 678 GGGCAGAGGGGACAGAGGGGACAGAGGGGACAGAGGGGACAGAGGGGACAGAG 737
Qy 823 GAGGACTCTAGAGAGGCGCTTCAG 849
Db 738 GAGCGCGGGTCAAGAGGACGTGAG 764

```

# RESULT 13 AAAS0254

ID AAAS0254 standard; DNA; 1926 BP.

AC AAAS0254;

DT 07-NOV-2000 (first entry)

DE Epstein Barr virus nuclear antigen 1 protein (EBNA1) DNA.

KM EBV; nuclear antigen 1; EBNA1; episome; transfection; selection;  
gene therapy; ds.

OS Epstein-barr virus.

PN WO20047778-A1.

PD 17-AUG-2000.

PE 11-FEB-2000; 2000WO-US03547.

PR 11-FEB-1999; 99US-0249585.

PA (PHAR-) PHARMACOEPIA INC.

PI Horlick RA, Chelsky D;

DR WPI; 2000-515062/46.

DR P-PSDB; AAY95856.

PT Stably transfecting eukaryotic cells with at least one episome for the  
production of a desired protein in vitro and for gene therapy -

PS Disclosure; Fig 2; 53pp; English.

XX The present sequence is that of DNA encoding the Epstein-Barr virus  
XX (EBV) nuclear antigen protein 1 (EBNA1, see AAY95856). EBNA1 is  
XX utilised in a novel method for obtaining a eukaryotic cell that is  
XX stably transfected with at least one episome. This method involves  
XX transfecting a eukaryotic cell with: (1) a first episome comprising  
XX an EBV origin of replication (oriP, see AAAS0253), a gene encoding a

CC first protein whose expression results in cell death and a  
CC selectable marker for eukaryotic cells; and (2) a second episome  
CC comprising an EBV oriP and a gene encoding a second protein, where  
CC expression of the second protein prohibits the occurrence of cell  
CC death resulting from expression of the first protein to produce  
CC doubly transfected cells which also express an antigen that  
CC promotes retention of the episomes by the cells. The doubly  
CC transfected cells are maintained under conditions in which the  
CC first and second proteins and the selectable marker are expressed,  
CC and the selective pressure specified by the marker is maintained.  
CC Under these conditions, only cells containing both episomes live.  
CC Preferably, EBNA1 is expressed from 1 of the episomes, and the  
CC protein of interest from the other episome. Either or both  
CC epitopes may further comprise a nucleic acid sequence encoding a  
CC protein desired to be expressed in the cell (e.g. a therapeutic  
CC protein), a nucleic acid encoding an RNA that is not intended to  
CC be translated (e.g. a therapeutic RNA), or a DNA sequence used as  
CC a tag for the cells. The method is applicable to cell culture or  
CC intact organisms, for gene therapy. It allows the rapid  
CC establishment of eukaryotic cells that stably and reliably express  
CC a gene of interest, using a novel method of selection, and  
CC maintenance of that selection without the need for exogenous  
CC selection factors, such as antibiotics.

Sequence 1926 BP; 487 A; 352 C; 872 G; 215 T; 0 other;

Query Match 4.4%; Score 63.8; DB 21; Length 1926;

Best Local Similarity 42.8%; Pred. No. 0.0051;

Matches 320; Conservative 0; Mismatches 427; Indels 0; Gaps 0;

```

Qy 103 GTGAGTCTCTGTCGACAGGTGCTCTGTGATGCTTACGCGACCCCAAGAG 162
Db 260 GTGGAACAGAGACAGAGACAGAGCGGGGACAGAGCGAGAGGGGACAGAG 319
Qy 163 GTGGAGGTACAGGGCTCTCAGGCTGCTTGGCAGAGAGCGGAGCCGAGCTG 222
Db 320 GAGGGGACAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 379
Qy 223 CTGCAGATGCTGAAGATCACCGACGACCCGAGCTGATCTGTCAGCTTGTGTC 282
Db 380 GAGAGGGGACAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 439
Qy 283 GGGCGGACCCCTGTGCGCTTCTCTCCGCGCTTACCGCGAGGGGCGCTGCGCCG 342
Db 440 GGGCAGAGACAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 499
Qy 343 CTGCAGAGAGAGCTTGGCGCGCGCTGCGCGAGCACTCGGTGCGCGCAACTGAG 402
Db 500 GAGGGGACAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 559
Qy 403 CGCGCGGCGCGAGCGGCTGGAAGCTTTGCTGGCGGACGAGAGCGCTTTGAGTTGC 462
Db 560 CAGAGAGGGGACAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 619
Qy 463 ATCTTACCCACACCCACCCGCTCCCGGATGAAAGACTGGAGTGAAGATGCG 522
Db 620 CAGAGAGGGGACAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 679
Qy 523 CTGCAAAATCTGAATGCGGCTCGGGGGCCCGGGGTGCGAGCGGAGGTGCTTCCGCC 582
Db 680 CAGAGAGGGGACAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 739
Qy 583 CCCTTGACGCCCCCGGTGCTCTCTGTGAGGTGAAGCCCGCGCCCGCACTT 642
Db 740 CAGAGAGGGGACAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 799
Qy 643 GCCCAGACTTTTCTGTTCAGGGGTACGCTGTAGTAATCGCGCGCTGAGCTTGAAGAC 702
Db 800 GAGGGGACAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 859
Qy 703 CAACAGAGTTCGCGGCTCTGTGGGTCTCAATGCGCAGAGTGGGCGCTCACTGAG 762
Db 860 CAGAGACAGAGAGGGGACAGAGGGGACAGAGAGAGAGAGAGAGAGAGAGAGAGAG 919

```





GenCore version 5.1.3  
Copyright (c) 1993 - 2003 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: February 3, 2003, 14:23:18 ; Search time 2216 Seconds  
(without alignments)  
10487.601 Million cell updates/sec

Title: US-09-763-748-1  
Perfect score: 1435  
Sequence: 1 ctgcgcggcgcgtgggaaccca.....gataataagatracacgcg 1435

Scoring table: IDENTITY\_NUC  
Gapop 10.0 , Gapext 1.0

Searched: 16154066 seqs, 8097743376 residues  
Total number of hits satisfying chosen parameters: 32308132

Minimum DB seq length: 0  
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Database :

EST: \*  
1: em\_estbda: \*  
2: em\_estbun: \*  
3: em\_estlin: \*  
4: em\_estnu: \*  
5: em\_estov: \*  
6: em\_estpl: \*  
7: em\_estro: \*  
8: em\_hic: \*  
9: gb\_est1: \*  
10: gb\_est2: \*  
11: gb\_hic: \*  
12: gb\_est3: \*  
13: gb\_est4: \*  
14: gb\_est5: \*  
15: em\_estfun: \*  
16: em\_estcom: \*  
17: gb\_gss: \*  
18: em\_gss\_hum: \*  
19: em\_gss\_inv: \*  
20: em\_gss\_pln: \*  
21: em\_gss\_vit: \*  
22: em\_gss\_fun: \*  
23: em\_gss\_mam: \*  
24: em\_gss\_mus: \*  
25: em\_gss\_other: \*  
26: em\_gss\_pro: \*  
27: em\_gss\_rtd: \*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

#### SUMMARIES

Result No.	Score	Query Length	DB ID	Description
1	765	53.3	1005	10 BE563501 601334982
2	725.8	50.6	1209	13 BE795807 601590842
3	651.2	45.4	976	13 BE193533 602947477
4	633	44.1	698	9 A1399660 602947477
5	629.8	43.9	810	12 BG327643 602426638
6	620.6	43.2	627	12 BF732310 602426638

Result No.	Score	Query Length	DB ID	Description
1	765	53.3	1005	10 BE563501 601334982
2	725.8	50.6	1209	13 BE795807 601590842
3	651.2	45.4	976	13 BE193533 602947477
4	633	44.1	698	9 A1399660 602947477
5	629.8	43.9	810	12 BG327643 602426638
6	620.6	43.2	627	12 BF732310 602426638

#### ALIGNMENTS

RESULT 1  
LOCUS BE563501 1005 bp mRNA EST 15-AUG-2000  
DEFINITION 601334982F1 NIH\_MGC\_39 Homo sapiens cDNA clone IMAGE:3688676 5',  
ACCESSION BE563501  
VERSION BE563501.1 GI:9807221  
KEYWORDS EST.  
SOURCE human.  
ORGANISM Homo sapiens  
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.  
REFERENCE 1 (bases 1 to 1005)  
AUTHORS NIH-MGC http://mgc.ncl.nih.gov/  
TITLES National Institutes of Health, Mammalian Gene Collection (MGC)  
JOURNAL Unpublished (1999)  
COMMENT Contact: Robert Strausberg, Ph.D.  
Email: cgabs-r@mail.nih.gov  
Tissue Procurement: ATCC  
CDNA Library Preparation: Ling Hong/Rubin Laboratory  
CDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)  
DNA Sequencing by: Incyte Genomics, Inc.  
Clone distribution: MGC clone distribution information can be  
found through the I.M.A.G.E. Consortium/LLNL at:  
http://image.llnl.gov  
Plate: LLCM382 row: d column: 21  
High quality sequence stop: 728.  
Location/Qualifiers 1..1005

FEATURES  
source

```

/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone="IMAGE:368676"
/clone_1ib="NIH MGC 39"
/tissue_type="adenocarcinoma"
/lab_note="DH10B (phage-resistant)"
/note="Organ: pancreas; Vector: pOTB7; Site:1: XhoI;
Site:2: EcoRI; cDNA made by oligo-dT priming.
Directionally cloned into EcoRI/XhoI sites using the
following 5' adaptor: GGCAACGG(G). Library constructed
by Ling Hong in the laboratory of Gerald M. Rubin
(University of California, Berkeley) using ZAP-cDNA
Synthesis Kit (Stratagene) and Superscript II RT (Life
Technologies)."
```

BASE COUNT 146 a 303 c 397 g 159 t

ORIGIN

Query Match 53.3%; Score 765; DB 10; Length 1005;

Best Local Similarity 93.9%; Pred. No. 7.4e-151;

Matches 872; Conservative 0; Mismatches 50; Indels 7; Gaps 7;

```

QY 11 TGGGAACCCAGAGCCCGCGAGGCGGCGAGAGTGATGAGCTGGGCAAAATGGGC 70
DB 1 TGGGAACCCAGAGCCCGCGAGGCGGCGAGAGTGATGAGCTGGGCAAAATGGGC 60
QY 71 ACGAAGAGTGGGTGGGCGAGCGCATCTCTTTGTGAGTCTCTGCTGACAAGGTGTCC 130
DB 61 ACGAAGAGTGGGTGGGCGAGCGCATCTCTTTGTGAGTCTCTGCTGACAAGGTGTCC 120
QY 131 TGTGGATGCTTACCGCGACCCCGACAGAGAGTGAGTGAAGGCTCTGACAGGCTG 190
DB 121 TGTGGATGCTTACCGCGACCCCGACAGAGAGTGAGTGAAGGCTCTGACAGGCTG 180
QY 191 CCTGGCAGAGAGCGGCGGAGAGCGGAGAGTGCTGAGATGCTGAGATCCACCGAGCG 250
DB 181 CTTGGCAGAGAGCGGCGGAGAGCGGAGAGTGCTGAGATGCTGAGATCCACCGAGCG 240
QY 251 ACCCGAGCTGATGCTGAGCTGCGATTCTGCGGCGGAGCGGCTGTGCGCTTCTCC 310
DB 241 ACCCGAGCTGATGCTGAGCTGCGATTCTGCGGCGGAGCGGCTGTGCGCTTCTCC 300
QY 311 GCGGCTTACCGCGAGGCGGCGCTGCGCGCGCTGCGAGAGAGCTTGGCGCGCTCG 370
DB 301 GCGGCTTACCGCGAGGCGGCGCTGCGCGCGCTGCGAGAGAGCTTGGCGCGCTCG 360
QY 371 CCAGAGCTCGAGGCGGCTGCGAACTGAGGTGCGCGCGCGCGCGAGCGGCTGGAAGCTT 430
DB 361 CCAGAGCTCGAGGCGGCTGCGAACTGAGGTGCGCGCGCGCGCGAGCGGCTGGAAGCTT 420
QY 431 TCGTGGCGAGCGAGAGCGCTGTTTGAAGTTCATCTTACCGCGAGCGCGAGCGGCTCC 490
DB 421 TCGTGGCGAGCGAGAGCGCTGTTTGAAGTTCATCTTACCGCGAGCGCGAGCGGCTCC 480
QY 491 GGGATGAAGAAGCTGCTGAGCTGAGAGAGTCCGTGGAATCTGAAGTCCGGTCCGGGG 550
DB 481 GGGATGAAGAAGCTGCTGAGCTGAGAGAGTCCGTGGAATCTGAAGTCCGGTCCGGGG 540
QY 551 CGCGGGGTGGCGAGCGGAGGTGCTTGGGCGCGCTTGAAGCGCGCGGCGCTCTCTGT 610
DB 541 CGCGGGGTGGCGAGCGGAGGTGCTTGGGCGCGCTTGAAGCGCGCGGCGCTCTCTGT 600
QY 611 CGAGAGTGAAGCGCGCGCGCGCGCGCGCGCGAGCTTCTTCTGTCGAGGCTCAGC 670
DB 601 CGAGAGTGAAGCGCGCGCGCGCGCGCGCGCGAGCTTCTTCTGTCGAGGCTCAGC 659
QY 671 CTGTAG-TGAATGCGCGCTGAGCTGAAGAGCAACAGAGCTTCTGCGCGCTCTGT-663 728
DB 660 CTGTAGTGAATGCGCGCTGAGCTGAAGAGCAACAGAGCTTCTGCGCGCTCTGT-719
QY 729 TCTCAATGCGCGAGAGTGGGCGGCTC-ACGAGAGAGAGCTGCGGCGGCTGCGGAGC 787
DB 720 TCTCAATGCGCGAGAGTGGGCGGCTC-ACGAGAGAGAGCTGCGGCGGCTGCGGAGC 779
```

```

QY 788 CGGCGCTGAGCTCGCTGAGCTTACGAGTACGAGCG-CGAGGAGCTGTACAGAGCGCTTC 846
DB 780 GGGCGCTGAGACTCGCTGAGTACGAGTACGAGCGCGCGAGGAGCTGTACAGAGCGCTTC 839
QY 847 CAGCTGCTGCGGCGCTTGTGTGAGGCGCGAGGCGCGCGCGCGAGCTTCTGAGGCTGTG 906
DB 840 CAGTGTGCGGCGGCTTGTGTGAGGCGCGAGGCGCGCGCGCGAGCTTCTGAGGCTGTG 897
QY 907 GAGGACTGAGAGAGAGAGAGAGCTCACAG 935
DB 898 GAGGCTCGGAGAGAGAGAGAGCTCACAGTCCACTG 926
```

#### RESULT 2

BE795807

LOCUS

DEFINITION

601590842P1 NIH\_MGC\_7 Homo sapiens cDNA clone IMAGE:3944795 5',

ACCESSION

VERSION

KEYWORDS

SOURCE

ORGANISM

REFERENCE

AUTHORS

TITLE

JOURNAL

COMMENT

Homo sapiens  
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.  
1 (bases 1 to 1209)  
NIH-MGC http://mgs.nci.nih.gov/  
National Institutes of Health, Mammalian Gene Collection (MGC)  
Unpublished (1999)  
Contact: Robert Strausberg, Ph.D.  
Email: cgapbs-remail.nih.gov  
Tissue Procurement: DCTD/DTF  
cDNA Library Preparation: Ling Hong/Rubin Laboratory  
DNA Sequencing by: Incyte Genomics, Inc.  
Clone distribution: MGC clone distribution/LINL at: image.lnl.gov  
Plate: LINC802 row: d column: 12  
High quality sequence stop: 750.  
Location/Qualifiers

#### FEATURES

source

```

1..1209
/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone="IMAGE:3944795"
/clone_1ib="NIH MGC 7"
/tissue_type="small_cell carcinoma"
/cell_line="MGC3"
/lab_host="DH10B (phage-resistant)"
/note="Organ: lung; Vector: pOTB7; Site:1: XhoI; Site:2:
EcoRI; cDNA made by oligo-dT priming. Directionally
cloned into EcoRI/XhoI sites using the following 5'
adaptor: GGCAACGG(G). Size-selected >500bp for average
insert size 1.8kb. Library constructed by Ling Hong in
the laboratory of Gerald M. Rubin (University of
California, Berkeley) using ZAP-cDNA synthesis kit
(Stratagene) and Superscript II RT (Life Technologies)."
```

#### BASE COUNT

218 a 379 c 434 g 178 t

#### ORIGIN

Query Match 50.6%; Score 725.8; DB 12; Length 1209;

Best Local Similarity 87.3%; Pred. No. 1.4e-142;

Matches 881; Conservative 0; Mismatches 107; Indels 21; Gaps 7;

```

QY 8 GCGTGGGAACCCAGAGCCCGCGCGAGGCGGCGAGAGTGATGAGCTGGGCAAAATG 67
DB 1 GCGTGGGAACCCAGAGCCCGCGCGAGGCGGCGAGAGTGATGAGCTGGGCAAAATG 60
QY 68 GCGAAGAGAGTGGGTGGGCGAGCGCATCTGTTGTGAGTCTCTGCTGACAAGGTG 127
DB 61 GCGAAGAGAGTGGGTGGGCGAGCGCATCTGTTGTGAGTCTCTGCTGACAAGGTG 120
QY 128 TCTGTGGATGCTTACCGCGAGCCCGCGAGAGAGTGAGTGAAGGCTGTGAGG 187
```

```

Db 121 TCCTGTGATGCTTACGCGACCCCGACGAAAGGTGGACGTGTACAGGCTTCGACG 180
Qy 188 CTGCTTGGAGAGAGCGGCGAGAGCCCGACGCTGACAGATGCTGAATCCACCGCA 247
Db 181 CTGCTTGGAGAGAGCGGCGAGAGCCCGACGCTGCTGACAGATGCTGAATCCACCGCA 240
Qy 248 GCGACCCGACGCTGATGCTGACAGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 307
Db 241 GCGACCCGACGCTGATGCTGACAGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 300
Qy 308 TCCGCGCTTACCGCGAGAGGCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 367
Db 301 TCCGCGCTTACCGCGAGAGGCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 360
Qy 368 TCCGCGACGACGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 427
Db 361 TCCGCGACGACGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 420
Qy 428 CTTTGTGCGGAGAGAGAGCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 487
Db 421 CTTTGTGCGGAGAGAGAGCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 480
Qy 488 TCCGCGATGAAAGACTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 547
Db 481 TCCGCGATGAAAGACTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 540
Qy 548 GCGCGCGGAGGAGAGAGAGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 607
Db 541 GCGCGCGGAGGAGAGAGAGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 600
Qy 608 TGTGGAAGTGAAGAGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCT 667
Db 601 TGTGGAAGTGAAGAGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCT 658
Qy 668 AGCTGTGATGATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 725
Db 659 CAGCTGTGATGATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 718
Qy 726 GGGTCTCAATAGGCGCA--GTTGGGCGCTCACTGAGAGAGGCTGCGGCGCTGCG 782
Db 719 GGGTCTCAATAGGCGCGCAAGTGGGCGCGCTCACTGAGAGAGGCTGCGGCGCTGCG 778
Qy 783 GGAACCGCGCTGAGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 835
Db 779 GGAACCGCGCG--GGACTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 837
Qy 836 AGCAGGCTTTCAGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 895
Db 838 AGCAGGCTTTCAGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 897
Qy 896 AGCGCTTGT-----GAGGCACTCGAGAGAGAGAGCTCAAGCTGCGCAAGACT 949
Db 898 AGCGCTTGTGTGAGGACCACTCGAGAGAGAGAGCTCAAGCTGCGCAAGACT 957
Qy 950 TGCTGGGCTGACCGATCCCAATGCGGCGCTGAGCTGAGACCGAGGCTGCTGCTGCTGCT 998
Db 958 GTTGGGCGCGCGAGAACACAGAGCGGCGCGCATTAACGCGGCGGCTGCTGCTGCTGCTGCT 1006

```

```

RESULT 3
B1193533
LOCUS 602947477F1 NIH_MGC_42 Homo sapiens cdna clone IMAGE:5090779 5',
DEFINITION mRNA sequence.
ACCESSION B1193533
VERSION B1193533.1 GI:14648540
KEYWORDS EST.
SOURCE human.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
REFERENCE 1 (bases 1 to 976)
AUTHORS NIH-MGC http://mgs.nci.nih.gov/.

```

```

TITLE National Institutes of Health, Mammalian Gene Collection (MGC)
JOURNAL Unpublished (1999)
COMMENT Contact: Robert Strausberg, Ph.D.
Email: cgaabs-remail.nih.gov
Tissue Procurement: ATCC
CDNA Library Preparation: Ling Hong/Rubin Laboratory
CDNA Library Arrayed by: The I.M.A.G.E. Consortium (LNL)
DNA Sequencing by: Incyte Genomics, Inc.
Clone distribution: MGC clone distribution information can be
found through the I.M.A.G.E. Consortium/LNL at:
http://image.llnl.gov
Plate: LNCMI853 row: 1 column: 20
High quality sequence scop: 745.
Location/Qualifiers
1. 976
/organism="Homo sapiens"
/ab_xref="taxon:9606"
/clone_1b="NIH_MGC_42"
/issue_type="epithelioid carcinoma cell line"
/lab_host="DH10B (phage-resistant)"
/notes="Organ: pancreas; Vector: pORF7; Site_1: XhoI;
Site_2: EcoRI; CDNA made by oligo-dT priming.
Directionally cloned into EcoRI/XhoI sites using the
following 5' adaptor: GGCACGAG(6). Size-selected >500bp
for average insert size 1.8kb. Library constructed by Ling
Hong in the laboratory of Gerald M. Rubin (University of
California, Berkeley) using ZAP-cDNA synthesis kit
(Stratagene) and Superscript II RT (Life Technologies).
Note: this is a NIH_MGC Library."
BASE COUNT 160 a 281 c 390 g 145 t
ORIGIN
Query Match 45.4%; Score 651.2; DB 13; Length 976;
Best Local Similarity 90.6%; Pred. No. 6.6e-127;
Matches 779; Conservative 0; Mismatches 58; Indels 23; Gaps 7;
Qy 1 CTGGCGGCGCTGGAAACCCAGCGCCCGCGAGCGCGCAAGAGTGAATGCGAGCTGG 60
Db 9 CTGGCGGCGCTGGAAACCCAGCGCCCGCGAGCGCGCGCAAGAGTGAATGCGAGCTGG 68
Qy 61 CAAATGAGGACGAAGAGTGGTGGGAGCGCACTACTGTTGTGAGATCTTCGCTGAC 120
Db 69 CAAATGAGGACGAAGAGTGGTGGGAGCGCACTACTGTTGTGAGATCTTCGCTGAC 128
Qy 121 AAGTGTCTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 180
Db 129 AAGTGTCTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 188
Qy 181 CTGAGGCTGCTTGGGAGAGAGCGGCGGAGCGCGGAGCGTGTGCGATGCTGAAGATC 240
Db 189 CTGAGGCTGCTTGGGAGAGAGCGGCGGAGCGCGGAGCGTGTGCGATGCTGAAGATC 248
Qy 241 CACCGAGCGACCCCGACGCTGATGTGACGCTGATCTTCGCGGCGCGACCCCTGTGC 300
Db 249 CACCGAGCGACCCCGACGCTGATGTGACGCTGATCTTCGCGGCGCGACCCCTGTGC 308
Qy 301 CGCTTCTCTCGCGCTTACCGCGAGGAGGCGCTGCGCGCGCTGCAAGAGACTTGGCG 360
Db 309 CGCTTCTCTCGCGCTTACCGCGAGGAGGCGCTGCGCGCGCTGCAAGAGACTTGGCG 368
Qy 361 GCGCGGCTGCGCGAGCACTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 420
Db 369 GCGCGGCTGCGCGAGCACTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 428
Qy 421 CTGAGCGCTTGTGCTGCGGAGAGAGCGCTGTTTGAAGTTGATCTTACCCAGACGCC 480
Db 429 CTGAGCGCTTGTGCTGCGGAGAGAGCGCTGTTTGAAGTTGATCTTACCCAGACGCC 488
Qy 481 GACCGGCTCGGAGATGAAGAACTGCGCTGAGCTGAGAGATGCGCTGCGAATCTGAAGTGC 540
Db 489 GACCGGCTCGGAGATGAAGAACTGCGCTGAGCTGAGAGATGCGCTGCGAATCTGAAGTGC 548

```

QY 541 GGCTCGGGGGCCCGGGGTGCGACGAGGAGTCTGCTTCGCCCCCTTGGACCCCCCGGTG 600  
 Db 549 GGCTCGGGGGCCCGGGGTGCGACGAGGAGGTGCTTCGCCCCCTTGGACCCCC--GGT 606  
 QY 601 CCTCTCTGTGCGAGGTGAAGCCGCCGCCGCCACCTGCCCCAAGATTCTTCTGTTT 660  
 Db 607 GCTCTCTGTGCGAGGTGAAGC-----GCCGCCGCCGCCACTGCGCAGATTCTCT 655  
 QY 661 CAGGGTCACTGTAGTGAATCGCGCTGAGCTGAAGAGCAACAAGAGTTTCCGCGCG 720  
 Db 656 CAGGTCACTGTAGTGAATCGCG--GTGAGCTTGAAGAGC--ACAGACGTTTCCGCGCG 711  
 QY 721 TCTGTGGGTCTCAAAATGCGCAGAGTGGGGCGCTCACTGACGAGGCTGCCGCGCTG 780  
 Db 712 TCTGTGGGTCTCAAAATGCGCG--AGTGGGGCGCTCAATG---GACGAGGTGCCGCGCG 766  
 QY 781 CCGGACCCCGCGCTGAGCTTCGCTGCGCTTACGAGTACGAGCGCG--AGGACTGTACGAGA 839  
 Db 767 CTGGGACCCGCGCTGAGCTTCGCTGCGCTTACGAGTACGAGGAGATGTCCGAGCA 826  
 QY 840 GGCTTCGAGCTGCTGCGCG 859  
 Db 827 GGCTTCGAGCTGATGGCG 846  
 RESULT 4  
 A1399660/c 698 bp mRNA linear EST 30-MAR-1999  
 LOCUS th28c09.x1 NCI\_CGAP\_Pr28 Homo sapiens cDNA clone IMAGE:2119600 3'  
 DEFINITION similar to SW:TRAD\_HUMAN Q15628 TUMOR NECROSIS FACTOR RECEPTOR TYPE  
 1 ASSOCIATED PROTEIN ; mRNA sequence.  
 ACCESSION A1399660  
 VERSION A1399660.1 GI:4242747  
 KEYWORDS EST.  
 SOURCE human.  
 ORGANISM Homo sapiens  
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 Mammalia; Eutheria; Primates; Catarrhini; Homidae; Homo.  
 REFERENCE 1 (bases 1 to 698)  
 NCI-CGAP http://www.ncbi.nlm.nih.gov/ncicgap.  
 TITLE National Cancer Institute, Cancer Genome Anatomy Project (CGAP),  
 Tumor Gene Index  
 JOURNAL Unpublished (1997)  
 COMMENT Contact: Robert Strausberg, Ph.D.  
 Email: cga@bbs-remail.nih.gov  
 Tissue procurement: Michael J. Brownstein, M.D., Ph.D., Michael R.  
 Emmer-Buck, M.D., Ph.D.  
 cDNA library preparation: M. Bento Soares, Ph.D.  
 cDNA library arrayed by: Greg Lennon, Ph.D.  
 DNA sequencing by: Washington University Genome Sequencing Center  
 Clone distribution: NCI-CGAP clone distribution information can be  
 found through the I.M.A.G.E. Consortium/ULNL at:  
 www.bio.linn.gov/bbtp/image/image.html  
 Insert length: 864 Std Error: 0.00  
 Seq primer: -40UP from Gibco  
 High quality sequence stop: 458.  
 Location/Qualifiers  
 1..698  
 /organism="Homo sapiens"  
 /db\_xref="taxon:9606"  
 /clone="IMAGE:2119600"  
 /clone\_lib="NCI\_CGAP\_Pr28"  
 /sex="male"  
 /dev\_stage="adult"  
 /lab\_host="DH10B"  
 /note="Organ: prostate; Vector: pT7T3D-Pac (Pharmacia)  
 with a modified polylinker; Plasmid DNA from the  
 normalized library NCI CGAP Pr22 was prepared, and ss  
 circles were made in vitro. Following HAP purification,  
 this DNA was used as tracer in a subtractive hybridization  
 reaction. The driver was PCR-amplified cDNAs from a pool  
 of 5,000 clones made from the same library (clonoids  
 985608-986759, 1101192-1101959, and 1217928-1220615).

BASE COUNT 150 a 202 c 200 g 146 t  
 ORIGIN  
 Query Match 44.1%; Score 633; DB 9; Length 698;  
 Best Local Similarity 96.9%; Pred. No. 4.1e-123;  
 Matches 656; Conservative 0; Mismatches 20; Indels 1; Gaps 1;  
 QY 760 CAGCAGAGCTGCGCGGCGCTGCGGAGCCGCGCTGAGCACTGCTGCGCTTACGAGTACAG 819  
 Db 698 CACAGAGGCTTCGGTCCCTCCGAGACCGCGCTGACTTGGCTGAGCTTACGAGTACAG 639  
 QY 820 CCGAGGAGCTGTACAGACAGGCC--TTCAGCTGTGCGGCGCTTCTGTACAGGCCAGAG 878  
 Db 638 CGCAGGAGCTGTACAGAACAGGCCCTTTCAAGTGTGTGCGGCGCTTCTGTACAGGCCAG 579  
 QY 879 CCGCGCGCGCAGCTGTGAGCGCGCTGTGTGAGGCACTGAGAGAAAGAGCTTACAGCTT 938  
 Db 578 CCGCGCGCGCAGCTGTGAGCGCGCTGTGTGAGGCACTGAGAGAAAGAGCTTACAGCTT 519  
 QY 939 GGCAGAGGACTTGTGCGGCGCTGACCGATCCCAATGCGGCGCTTGTAGACAGAGGTCG 998  
 Db 518 GGCAGAGGACTTGTGCGGCGCTGACCGATCCCAATGCGGCGCTTGTAGACAGAGGTCG 459  
 QY 999 AGCCAGCTTTTGGAGAACCTGAGATGCGCTTGTGCGGCTTGTGAGTGTGAGAACCC 1058  
 Db 458 AGCCAGCTTTTGGAGAACCTGAGATGCGCTTGTGCGGCTTGTGAGTGTGAGAACCC 399  
 QY 1059 CTGTTCATCCACGGGACCTTGAATCTCACTTGGCTTATCTGTGAGACTTGTGCGGCGAG 1118  
 Db 398 CTGTTCATCCACGGGACCTTGAATCTCACTTGGCTTATCTGTGAGACTTGTGCGGCGAG 339  
 QY 1119 AGTTGATTCGCTTCCCGAGGAGCCAGACCACTGGGGGGTGCATCATTTGGGGATTCGCTC 1178  
 Db 338 AGTTGATTCGCTTCCCGAGGAGCCAGACCACTGGGGGGTGCATCATTTGGGGATTCGCTC 279  
 QY 1179 AGGTACTTGTATAGAGTGTGGGGTGGGGGAGCTTGTGAGATCAGCTCAGCTTCT 1238  
 Db 278 AGGTACTTGTATAGAGTGTGGGGTGGGGGAGCTTGTGAGATCAGCTCAGCTTCT 219  
 QY 1239 CCATCCCAAGGCGGGCTTACAGCCAGCCCTTACAGTTTCACTCATGAAGACACTTGA 1298  
 Db 218 CCATCCCAAGGCGGGCTTACAGCCAGCCCTTACAGTTTCACTCATGAAGACACTTGA 159  
 QY 1299 TCTTTGGTGTCTGTGAGCTTCAATCCGCGGTGTGAGATACAGTGAAGTAAACAGAGA 1358  
 Db 158 TCTTTGGTGTCTGTGAGCTTCAATCCGCGGTGTGAGATACAGTGAAGTAAACAGAGA 99  
 QY 1359 ATCAATCTTGTGCTGCGCGCCAGCTCACTCAAGCGTGGAGCCGAGATGTAAAGATAT 1418  
 Db 98 ATCAATCTTGTGCTGCGCGCCAGCTCACTCAAGCGTGGAGCCGAGATGTAAAGATAT 39  
 QY 1419 AATTAAGTATTAACAGG 1435  
 Db 38 AATTAAGTATTAACCGG 22  
 RESULT 5  
 BG327643 810 bp mRNA linear EST 27-FEB-2001  
 LOCUS 60242638F1 NIH\_MGC\_14 Homo sapiens cDNA clone IMAGE:4564736 5',  
 DEFINITION mRNA sequence.  
 ACCESSION BG327643  
 VERSION BG327643.1 GI:13134081  
 KEYWORDS EST.  
 SOURCE human.  
 ORGANISM Homo sapiens  
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 Mammalia; Eutheria; Primates; Catarrhini; Homidae; Homo.  
 REFERENCE 1 (bases 1 to 810)  
 NIH-MGC http://mgc.nci.nih.gov/.  
 AUTHORS NIH-MGC  
 TITLE National Institutes of Health, Mammalian Gene Collection (MGC)  
 JOURNAL Unpublished (1999)



COMMENT Contact: Robert Strausberg, Ph.D.  
Email: cgsabs-r@mail.nih.gov  
Tissue Procurement: DCTD/BTP  
CDNA Library Preparation: Ling Hong/Rubin Laboratory  
CDNA Library Arrayed by: The I.M.A.G.E. Consortium (LNL)  
DNA Sequencing by: Inceye Genomics, Inc.  
Clone distribution: MGC clone distribution information can be  
found through the I.M.A.G.E. Consortium/LNL at:  
http://image.llnl.gov  
plate: L1CMI280 row: k column: 09  
High quality sequence stop: 622.  
Location/Qualifiers

FEATURES  
source  
1. 810  
/organism="Homo sapiens"  
/db\_xref="taxon:9606"  
/clone="IMAGE:4564736"  
/clone\_lib="NIH\_MGC\_14"  
/tissue\_type="renal cell adenocarcinoma"  
/lab\_host="DH10B (phage-resistant)"  
/note="Organ: kidney; Vector: pOTB7; Site 1: XhoI; Site 2:  
EcoRI; cDNA made by oligo-dT priming. Directionally  
cloned into EcoRI/XhoI sites using the following 5'  
adapter: GGCAAGAG(G). Size selected >500bp for average  
insert size 1.8kb. Library constructed by Ling Hong in  
the laboratory of Gerald M. Rubin (University of  
California, Berkeley) using ZAP-cDNA synthesis kit  
(Stratagene) and Superscript II RT (Life Technologies)."

BASE COUNT 140 a 256 c 308 g 106 t

ORIGIN

Query Match 43.9%; Score 629.8; DB 12; Length 810;  
Best Local Similarity 94.9%; Pred. No. 2e-122;  
Matches 682; Conservative 0; Mismatches 33; Indels 4; Gaps 3;

1 CTGGCGGCGGTGGGAACCCAGGCCCCCGGCGGCGGAGGTTGAGTGGAGCTGGG 60  
14 CTGGCGGCGGTGGGAACCCAGGCCCCCGGCGGCGGAGGTTGAGTGGAGCTGGG 73  
61 CAAATATGGGACGGAAGTGGTGGGAGCGGATACCTGTTGAGTGGAGTGGAGCTGG 120  
74 CAAATATGGGACGGAAGTGGTGGGAGCGGATACCTGTTGAGTGGAGTGGAGCTGG 133  
121 AAGGTGTCCTGTCGATGATGCTTACGCGGACCCCGAGAGAGTGGAGTGGAGCTGG 180  
134 AAGGTGTCCTGTCGATGATGCTTACGCGGACCCCGAGAGAGTGGAGTGGAGCTGG 193  
181 CTGCAAGCTGCTTGGGAGAGAGCGGCGGAGGCGGAGCTGTCGACAGTCTGAAGATC 240  
194 CTGCAAGCTGCTTGGGAGAGAGCGGCGGAGGCGGAGCTGTCGACAGTCTGAAGATC 253  
241 CACCGGAGCGACCGGAGCTGATGTCGACGTCGATTCGCGGCGGAGGCTGTCGCGC 300  
254 CACCGGAGCGACCGGAGCTGATGTCGACGTCGATTCGCGGCGGAGGCTGTCGCGC 313  
301 CGCTTCTCCGCGCTTACCGGAGGAGGCGGCGGCGGCTGTCGAGAGAGGCTGGCG 360  
314 CGCTTCTCCGCGCTTACCGGAGGAGGCGGCGGCGGCTGTCGAGAGAGGCTGGCG 373  
361 GCGGCGCTCGCGGAGACTCGGTGCGGCTGCACTGAGTGGCGCGCGCGGAGCGG 420  
374 GCGGCGCTCGCGGAGACTCGGTGCGGCTGCACTGAGTGGCGCGCGCGGAGCGG 433  
421 CTGGAACCTTGTGTCGCGGAGAGAGCGGCTGTTGAGTGGAGTGGAGTGGAGTGG 480  
434 CTGGAACCTTGTGTCGCGGAGAGAGCGGCTGTTGAGTGGAGTGGAGTGGAGTGG 493  
481 GACCGGCTCGCGGAGTGAAGAACTGCTGAGCTGAGAGTGGCTGCAAAATCTGAATGC 540  
494 GACCGGCTCGCGGAGTGAAGAACTGCTGAGCTGAGAGTGGCTGCAAAATCTGAATGC 553  
541 GCGCTCGGAGGCGCGGAGTGGAGCGGAGAGGTCG-CTTGGCGCGCTTGGAGCGCGCGG 599  
554 GCGCTCGGAGGCGCGGAGTGGAGCGGAGAGGTCGAAACGAGCGCGCTTGGAGCGCGG 613

600 GCCTCTCTGTCGAGAGTGAAGCGCGCGCGCGCGCCACCTGCGCAGACTTTCTGTT 659  
614 GCCTCTCTGTCGAGAGTGAAGCGCGCGCGCGCGCCACCTGCGCAGACTTTCTGTT 673  
660 CCAAGGTCAACCTGTA-ATGAATCGGCGCGCTGAG-CTTGAAGAGCAACAGACGTTG 715  
674 CAAGGCGCACCGCGACCGCAATTCGGCGCGCTGAGCGCTGAAGAGCAACAGACGATCG 732

RESULT 6  
BF732310  
LOCUS  
DEFINITION  
BF732310 627 bp mRNA linear EST 30-MAR-2001  
nae09a08.x1 NCI CGAP OV18 Homo sapiens cDNA clone IMAGE:3434774.3,  
similar to SW-TRAD HUMAN O15628 TUMOR NECROSIS FACTOR RECEPTOR TYPE  
1 ASSOCIATED DEATH DOMAIN PROTEIN; contains element TARI repetitive  
element; mRNA sequence.  
BF732310  
GI:12057385

ACCESSION  
VERSION  
KEYWORDS  
SOURCE  
ORGANISM  
Homo sapiens  
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.  
1 (bases 1 to 627)  
NCI-CGAP http://www.ncbi.nlm.nih.gov/ncicgap.  
National Cancer Institute, Cancer Genome Anatomy Project (CGAP),  
Tumor Gene Index  
Unpublished (1997)  
CONTACT: Robert Strausberg, Ph.D.  
Email: cgsabs-r@mail.nih.gov  
Tissue Procurement: Christopher A. Moskaluk, M.D., Ph.D., Michael  
R. Emmert-Buck, M.D., Ph.D. CDNA Library Preparation: M. Bento  
Soares, Ph.D. CDNA Library Arrayed by: Christa Frange, The  
I.M.A.G.E. Consortium DNA Sequencing by: Washington University  
Genome Sequencing Center  
Clone distribution: NCI-CGAP clone distribution information can be  
found through the I.M.A.G.E. Consortium/LNL, send email to:  
info@image.llnl.gov  
Seq primer: -400p from Gibco  
High quality sequence stop: 445.  
Location/Qualifiers  
1. 627  
/organism="Homo sapiens"  
/db\_xref="taxon:9606"  
/clone="IMAGE:3434774"  
/clone\_lib="NCI CGAP OV18"  
/tissue\_type="fibrotheoma"  
/lab\_host="DH10B (phage-resistant)"  
/note="Organ: ovary; Vector: pRT3D-Pac (pharmacia) with a  
modified polylinker; Site 1: Not I; Site 2: Eco RI; 1st  
strand cDNA was primed with a Not I - oligo(dT) primer (5'  
TGTTACATCTGAAGTGGAGCGCGCGGCGGACATTTTCTTTTCTTTT 3');  
double-stranded cDNA was ligated to Eco RI adaptors  
(pharmacia), digested with Not I and cloned into the Not  
I and Eco RI sites of the modified pRT3 vector. Library  
went through one round of normalization, and was  
constructed by Bento Soares and M. Fatima Bonaldo."

BASE COUNT 91 a 206 c 231 g 99 t

ORIGIN

Query Match 43.2%; Score 620.6; DB 12; Length 627;  
Best Local Similarity 99.4%; Pred. No. 1.6e-120;  
Matches 623; Conservative 0; Mismatches 4; Indels 0; Gaps 0;

358 GCGGCGGCTCGCGGAGACTCGGTGCGGCTGCACTGAGTGGCGCGCGCGCGAG 417  
Db 1 GCGGCGGCTCGCGGAGACTCGGTGCGGCTGCACTGAGTGGCGCGCGCGCGAG 60  
418 CCGGTGAGCCCTTGTGTCGCGGAGAGAGCGCTGTTGAGTGGATCTTGAAGCCAGAG 477  
Db 61 CCGGTGAGCCCTTGTGTCGCGGAGAGAGCGCTGTTGAGTGGATCTTGAAGCCAGAG 120

Qy 478 CCGGACCGGCTCCGGGATGAAGAACTGAGCTGAGCTGAGAGATGCGTTCGGAATCTGAAG 537  
Db 121 CCGGACCGGCTCCGGGATGAAGAACTGAGCTGAGAGATGCGTTCGGAATCTGAAG 180  
Qy 538 TCGGCTCGGG 597  
Db 181 TCGGCTCGGG 240  
Qy 598 GGGGCTCTCTGTCGGAGGTGAAGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 657  
Db 241 GGGGCTCTCTGTCGGAGGTGAAGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 300  
Qy 658 TTCCAGGGTCAAGCTTGTAGTGAATCGCGCGCTGAGCGCTGAGAGACCAAGAGCTTGGCG 717  
Db 301 TTCCAGGGTCAAGCTTGTAGTGAATCGCGCGCTGAGCGCTGAGAGACCAAGAGCTTGGCG 360  
Qy 718 CGCTCTGTGGGTCTCAAAATGGGCGAAGGTGGGGCGCTCACTGAGAGAGGCTTGGCGGGG 777  
Db 361 CGCTCTGTGGGTCTCAAAATGGGCGAAGGTGGGGCGCTCACTGAGAGAGGCTTGGCGGGG 420  
Qy 778 CTGGCGGACCGCGCGCTGAGCTGCTGGCTTACGAGTACGAGCGCGAGGAGCTGTACGAG 837  
Db 421 CTGGCGGACCGCGCGCTGAGCTGCTGGCTTACGAGTACGAGCGCGAGGAGCTGTACGAG 480  
Qy 838 CAGGCTTTCAGCTGCTGGCGGCTTCTGTCAGAGCGCGCGCGCGCGCGCGCGCGCGCGCG 897  
Db 481 CAGGCTTTCAGCTGCTGGCGGCTTCTGTCAGAGCGCGCGCGCGCGCGCGCGCGCGCGCG 540  
Qy 898 CGCTGTGTGGAGCACTCGAGAGAAAGAGCTTACAGGCTTGGCAGAGAGAGCTTGGCGGG 957  
Db 541 CGCTGTGTGGAGCACTCGAGAGAAAGAGCTTACAGGCTTGGCAGAGAGAGCTTGGCGGG 600  
Qy 958 CTGACCGATCCCAATGGCGGCTGGCC 984  
Db 601 CTGACCGATCCCAATGGCTGGCGGCTG 627

RESULT 7  
BM973321/c 660 bp mRNA linear EST 21-MAR-2002  
LOCUS BM973321  
DEFINITION UT-CF-EC1-abx-e-21-0-UI s1 UT-CF-EC1 Homo sapiens cDNA clone

ACCESSION BM973321  
VERSION BM973321.1 GI:19590912  
KEYWORDS EST.  
SOURCE human.

## ORGANISM

Homo sapiens  
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.

REFERENCE 1 (bases 1 to 660)  
AUTHORS Ronaldo M.F., Lennon G. and Soares M.B.  
TITLE Normalization and subtraction: two approaches to facilitate gene  
discovery

JOURNAL Genome Res. 6 (9), 791-806 (1996)  
MEDLINE 97044477  
COMMENT Contact: McCray, PB  
McCray Lab  
University of Iowa  
2024 University of Iowa Med Labs, Iowa City, IA 52242, USA  
Tel: 319 356 4866  
Fax: 319 356 7171

Email: paul-mccray@uiowa.edu  
Tissue Procurement: Dr. M. J. Welsh, University of Iowa  
cDNA Library preparation: Dr. M. Bento Soares, University of Iowa  
cDNA Library Arrayed by: Dr. M. Bento Soares, University of Iowa  
DNA Sequencing by: Dr. M. Bento Soares, University of Iowa  
Clone Distribution: Researchers may obtain clones from Research  
Genetics (www.resgen.com).

The following repetitive elements were found in this cDNA  
sequence: 77-145, >LINE2  
Seq primer: M13 FORWARD  
POLYA=Yes.

FEATURES  
Location/Qualifiers

source  
1. 660  
/organism="Homo sapiens"  
/db\_xref="taxon:9606"  
/clone="UI-CF-EC1-abx-e-21-0-UI"  
/clone\_lib="UI-CF-EC1"  
/tissue\_type="Lung"  
/dev\_stage="Adult and Fetal"  
/lab\_host="DH10B (Life Technologies) (T1 phage resistant)"  
/note="Organ: Lung; Vector: pT73-Pac (Pharmacia) with a  
modified polylinker; Site 1: EcoR I; Site 2: Not I;  
UI-CF-EC1 is a normalized cDNA library containing the  
following tissue(s): Normal lung from adult and from fetal  
day 64, day 87, week 19 and week 42. The library was  
constructed according to Bonaldo, Lennon and Soares,  
Genome Research, 6:791-806, 1996. First strand cDNA  
synthesis was primed with an oligo-dT primer containing a  
Not I site. Double stranded cDNA was ligated to an EcoR I  
adaptor, digested with Not I, and cloned directionally  
into pT73-Pac vector. The oligonucleotide used to prime  
the synthesis of first-strand cDNA contains a library tag  
sequence that is located between the Not I site and the  
(dT)18 tail. The sequence tag for this library is  
AAGTCTTAC.  
TAG\_L1B=UI-CF-EC1  
TAG\_TISSUE=Normal Lung Epithelial Cells Tissue nos 369-371  
and 380-383  
TAG\_SEQ=AAAGTCTTAC"  
BASE COUNT 139 a 191 c 184 g 146 t  
ORIGIN  
Query Match 43.0%; Score 616.8; DB 14; Length 660;  
Best Local Similarity 99.5%; Pred. No. 1e-119;  
Matches 629; Conservative 0; Mismatches 2; Indels 1; Gaps 1;  
Qy 804 GGCCTACAGATACAGAGCGCGGAGGAGCTGAGAGAGGCTTCCAGCTGCTGGCGGCTT 863  
Db 660 GGCCTACAGATACAGAGCGCGGAGGAGCTGAGAGAGGCTTCCAGCTGCTGGCGGCTT 601  
Qy 864 CGTGAGGCGGAGGCG 923  
Db 600 CGTGAGGCGGAGGCG 542  
Qy 924 CGAGCTACAGAGCTTGGAGAGAGCTTGGAGAGAGCTTGGAGAGAGCTTGGAGAGAG 983  
Db 541 CGAGCTACAGAGCTTGGAGAGAGCTTGGAGAGAGCTTGGAGAGAGCTTGGAGAGAG 482  
Qy 984 CTAGACCAAGGGTGCAGGAGGAGCTTGGAGAGAGCTTGGAGAGAGCTTGGAGAGAG 1043  
Db 481 CTAGACCAAGGGTGCAGGAGGAGCTTGGAGAGAGCTTGGAGAGAGCTTGGAGAGAG 422  
Qy 1044 GCTATTGCTGAACCCCTGTCACATCCAGAGGAGCCCTGAAGACTCCATTTGGCTGCTG 1103  
Db 421 GCTATTGCTGAACCCCTGTCACATCCAGAGGAGCCCTGAAGACTCCATTTGGCTGCTG 362  
Qy 1104 GACCTGTGGGGCAGAGTGTATGCTTCCAGAGGAGCAGACCACTGGGGGTGATCAT 1163  
Db 361 GACCTGTGGGGCAGAGTGTATGCTTCCAGAGGAGCAGACCACTGGGGGTGATCAT 302  
Qy 1164 TGGGATTTCTGCTCAGGATCTTTATATGAGTGTGGGGTGGGGGAGATTTGCTTGGAGA 1223  
Db 301 TGGGATTTCTGCTCAGGATCTTTATATGAGTGTGGGGTGGGGGAGATTTGCTTGGAGA 242  
Qy 1224 TCAGGCTACCTTCCATCCATCCAGAGGAGGCTTACAGGAGGCTTACAGTTCAT 1283  
Db 241 TCAGGCTACCTTCCATCCATCCAGAGGAGGCTTACAGGAGGCTTACAGTTCAT 182  
Qy 1284 CATGAAGACCTTGAATCTTGTGTCTCTGAGCTTCACTTCCTGGTGTCTCAATCTGCAG 1343  
Db 181 CATGAAGACCTTGAATCTTGTGTCTCTGAGCTTCACTTCCTGGTGTCTCAATCTGCAG 122  
Qy 1344 TGAAGTAAACAGGAATCAATCTTGCCTGCCAGCTCAGCTCAGCTCAGCTCAGCTCAGCT 1403  
Db 121 TGAAGTAAACAGGAATCAATCTTGCCTGCCAGCTCAGCTCAGCTCAGCTCAGCTCAGCT 62

QY 1404 ATGTTAGCAATGATTAATAAGTATAACACGG 1435  
 Db 61 ATGTTAGCAATGATTAATAAGTATAACACGG 30

RESULT 8  
 BG342078 788 bp mRNA linear EST 27-FEB-2001  
 LOCUS 602463023F1 NIH\_MGC\_48 Homo sapiens cDNA clone IMAGE:4575941 5',  
 DEFINITION mRNA sequence.  
 ACCESSION BG342078 GI:13148504  
 VERSION BG342078.1  
 KEYWORDS EST.  
 SOURCE human.  
 ORGANISM Homo sapiens  
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.  
 REFERENCE 1 (bases 1 to 788)  
 NIH-MGC http://mgs.nci.nih.gov/.  
 AUTHORS National Institutes of Health, Mammalian Gene Collection (MGC)  
 JOURNAL Unpublished (1999)  
 COMMENT Contact: Robert Strausberg, Ph.D.  
 Email: cgabs-remail.nih.gov  
 Tissue Procurement: Louis M. Staudt, M.D., Ph.D.  
 cDNA Library Preparation: Ling Hong/Rubin Laboratory  
 DNA Sequencing by: Incyte Genomics, Inc.  
 Clone distribution: MGC clone distribution information can be  
 found through the I.M.A.G.E. Consortium/LLNL at:  
 http://image.llnl.gov  
 Plate: LLCM1287 row: n column: 06  
 High quality sequence stop: 627.  
 Location/Qualifiers  
 1..788  
 /organism="Homo sapiens"  
 /db\_xref="taxon:9606"  
 /clone\_image="IMAGE:4575941"  
 /clone\_1ib="NIH MGC 48"  
 /tissue\_type="Primary B-cells from tonsils (cell line)"  
 /lab\_host="DH10B (phage-resistant)"  
 /note="Organ: B-cells; Vector: pOTB7; Site: 1: XhoI;  
 Site 2: EcoRI; cDNA made by oligo-dT priming.  
 Directionally cloned into EcoRI/XhoI sites using the  
 following 5' adaptor: GGCAAGG(G). Size-selected >500bp  
 for average insert size 1.8kb. Library constructed by Ling  
 Hong in the laboratory of Gerald M. Rubin (University of  
 California, Berkeley) using ZAP-cDNA synthesis kit  
 (Stratagene) and Superscript II RT (Life Technologies).  
 Note: this is a NIH\_MGC Library."

BASE COUNT 121 a 234 c 305 g 128 t

Query Match 42.8%; Score 613.8; DB 12; Length 788;  
 Best Local Similarity 93.2%; Pred. No. 4.6e-119;  
 Matches 709; Conservative 0; Mismatches 42; Indels 10; Gaps 6;

QY 1 CTGCGCGGCGTGGGAACCCAGGCCCCCGAGGGGCGGAGAGTGAATGGACGCTGGG 60  
 Db 14 CTGCGCGGCGTGGGAACCCAGGCCCCCGAGGGGCGGAGAGTGAATGGACGCTGGG 72

QY 61 CAAATATGGGACGAGAGAGTGGGAGGCGCATACCTGTTGGAGTCTCTGCTGGAC 120  
 Db 73 CAAATATGGGACGAGAGAGTGGGAGGCGCATACCTGTTGGAGTCTCTGCTGGAC 132

QY 121 AAGGTGTCTCTGTGAGTGTCTACGCGACCCCGAGAGAGAGTGGAGTGAACGAGCT 180  
 Db 133 AAGGTGTCTCTGTGAGTGTCTACGCGACCCCGAGAGAGAGTGGAGTGAACGAGCT 192

QY 181 CTGAGAGCTGCTTGGAGAGAGAGCGGGGAGCGCGGAGCTGCTGCGAGTCTAAAGATC 240  
 Db 193 CTGAGAGCTGCTTGGAGAGAGAGCGGGGAGCGCGGAGCTGCTGCGAGTCTAAAGATC 252

QY 241 CACCGAGCGACCCCGAGAGTGAATCTGACAGCTGCAATTTCTGCGGGGCGGACCCCTGTGAC 300  
 Db 253 CACCGAGCGACCCCGAGAGTGAATCTGACAGCTGCAATTTCTGCGGGGCGGACCCCTGTGAC 312

QY 301 CGCTTCTCTCCGCGCTACCGCGAGGGGGCGCTGCGCGCGCGCTGCAAGAGACCTGGCG 360  
 Db 313 CGCTTCTCTCCGCGCTACCGCGAGGGGGCGCTGCGCGCGCGCTGCAAGAGACCTGGCG 372

QY 361 GCGCGCTGCGCGCGAGAGTGAATCTGACAGCTGCAATTTCTGCGGGGCGGACCCCTGTGAC 420  
 Db 373 GCGCGCTGCGCGCGAGAGTGAATCTGACAGCTGCAATTTCTGCGGGGCGGACCCCTGTGAC 432

QY 421 CTGAGAGCTTGTGCTGGGAGAGAGAGCGCTGTTTGAAGTTGATCTTGAAGCCAGAGCCC 480  
 Db 433 CTGAGAGCTTGTGCTGGGAGAGAGAGCGCTGTTTGAAGTTGATCTTGAAGCCAGAGCCC 492

QY 481 GACCGGCTCCGGAGATGAAGAACTGCTGAGCTGAGAGATGCTGCGAATCTGAAGTGC 540  
 Db 493 GACCGGCTCCGGAGATGAAGAACTGCTGAGAGATGCTGCGAATCTGAAGTGC 552

QY 541 GCGTGGGGGCGCGGGGGTGGCG-ACGGGAGAGTGGCTTGGGCCCC-TTGACAGCCCCGG 598  
 Db 553 GCGTGGGGGCGCGGGGGTGGCG-ACGGGAGAGTGGCTTGGGCCCC-TTGACAGCCCCGG 612

QY 599 TGCCCTCTCTGTGAGAGTGAAGCCGCGCGCGCGCCAGCCTTCTGT 658  
 Db 613 TGCCCTCTCTGTGAGAGTGAAGCCGCGCGCGCGCCAGCCTTCTGT 669

QY 659 TCAGAGGTCAGCTGTGATGATTCGCGCGCTGAGCTGAAGACCAACAGACCTTGGCG 718  
 Db 670 TCAGAGGTCAGCTGTGATGATTCGCGCGCTGAGCTGAAGACCAACAGACCTTGGCG 728

QY 719 GCTGTGGGTCTCAATGCGGAGTGGGGGCGCTACACG 759  
 Db 729 -TCTGTGGGTCTCAATGCGGAGTGGGGGCGCTACACG 766

RESULT 9  
 BG746223 987 bp mRNA linear EST 15-MAY-2001  
 LOCUS 602723748T1 NIH\_MGC\_113 Homo sapiens cDNA clone IMAGE:4850170 3',  
 DEFINITION mRNA sequence.  
 ACCESSION BG746223  
 VERSION BG746223.1 GI:14056876  
 KEYWORDS EST.  
 SOURCE human.  
 ORGANISM Homo sapiens  
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.  
 REFERENCE 1 (bases 1 to 987)  
 NIH-MGC http://mgs.nci.nih.gov/.  
 AUTHORS National Institutes of Health, Mammalian Gene Collection (MGC)  
 JOURNAL Unpublished (1999)  
 COMMENT Contact: Robert Strausberg, Ph.D.  
 Email: cgabs-remail.nih.gov  
 Tissue Procurement: Dr. Mark Watson  
 cDNA Library Preparation: Ling Hong/Rubin Laboratory  
 DNA Sequencing by: Incyte Genomics, Inc.  
 Clone distribution: MGC clone distribution information can be  
 found through the I.M.A.G.E. Consortium/LLNL at:  
 http://image.llnl.gov  
 Plate: LLCM1690 row: p column: 11  
 High quality sequence start: 27  
 High quality sequence stop: 788.  
 Location/Qualifiers  
 1..987  
 /organism="Homo sapiens"  
 /db\_xref="taxon:9606"  
 /clone\_image="IMAGE:4850170"  
 /clone\_1ib="NIH MGC 113"  
 /lab\_host="DH10B (phage-resistant)"  
 /note="Organ: spleen; Vector: pOTB7; Site: 1: XhoI; Site: 2:

EscoRI, cDNA made by oligo-dT priming. Directionally cloned into EscoRI/XhoI sites using the following 5' adaptor: GGACAG(G). Library constructed by Ling Hong in the laboratory of Gerald M. Rubin (University of California, Berkeley) using ZAP-cDNA synthesis kit (Stratagene) and Superscript II RT (Life Technologies). Note: this is a NIH-MGC Library."

BASE COUNT 209 a 301 c 304 g 173 t  
ORIGIN

Query Match 42.7%; Score 613.4; DB 12; Length 987;  
Best Local Similarity 98.4%; Pred. No. 5.9e-119;  
Matches 827; Conservative 0; Mismatches 86; Indels 22; Gaps 14;

```
Qy 488 TCCGGGATGAGAACTGCGTGTAGTGTAGAGATCGCTCGGAATCTGAAATGTGCGGCTCG 547
Db 935 TTCGGGATGAGAACTGCGTGTAGTGTAGAGATCGCTCGGAATCTGAAATGTGCGGCTCG 876
Qy 548 GGG-CCCGGGGTGGCGAGCGGGAGGTGCGTTCGGCCCTTGCAGC-CCCGGTGCTTC 605
Db 875 GGGTCCCGGGGTGGCGAGCGGGAGGTGCGTTCGGCCCTTGCAGCCTTCCGGTCCCTC 816
Qy 606 TCTGCGAGAGGAGAGCCGCCGCCGCCGACCTGCGCAGACTT-TCGTGTCAGG 664
Db 815 TCTGCGAGAGCTGAGACCCCGCTGCCCGACCTGCGCAGACTTCTTGTTCAGG 756
Qy 665 GTACACCT-GTAGTGAATCGCGCTGAGCTGAGGACCAACAGACTTCCGCGCTC- 722
Db 755 GTACACCTGTAGTGAATCGCGCTGAGGCTGAGGACCTTCAAGCTGTGCGGCTCT 696
Qy 723 TGTGGGTCTCAATGCGCGCAGAGTGGGGCGCTCA--CTGACGAGAGCTGCGGGCG-CT 779
Db 695 TCTGGGTCTCAATGCGCGCAGAGTGGGGCGCTCACTGTGAGACGAGGCTGCGGGCGACT 636
Qy 780 GCGGAGCCCGGGGCTGGAATCGTGGGCTTACAGATACAGCGCGAGGACTGTAGAGCA 839
Db 635 GCGGAGCCCGGGGCTGGAATCGTGGGCTTACAGATACAGCGCGAGGACTGTAGAGCA 576
Qy 840 GGCC--TTCCAGCTGCTGCGCGCTTCTGCAAGCGCGAGCGCGCCGAC-CTGCA 896
Db 575 GGCCATTCAGAGATGCGCGCGCTTCTGCAAGCGCGAGCGCGCCGCTGCTGCA 516
Qy 897 GCGCTGTGTGAGG----CACTCGAGAGAACGAGCTCAACGAGCTGCGAGAGCACTTG 951
Db 515 GCGCTGTGTGAGGCACTCGAGAGAGAACGAGCTCAACGAGCTGCGAGAGCACTTG 456
Qy 952 CTGGGCTGAGCCGATCCGATGGCGGCTGAGTACAGGGGTGACGAGCTTTG 1011
Db 455 CTGGGCTGAGCCGATCCGATGGCGGCTGAGTACAGGGGTGACGAGCTTTG 396
Qy 1012 AGAACCTGATGGCTTAAAGGTTCTTCTGCGG-CTATTGTGAACCCCTGTCCATCCAC 1070
Db 395 AGAACCTGATGGCTTAAAGGTTCTTCTGCGGCTTATGGCTGAACCCCTGTCCATCCAC 336
Qy 1071 GGGACCTGAAAATCCACTTGGCTTATCTGTGAGACTT-GCTGGGGCAGAGTTGATGCC 1129
Db 335 GGGAACTGAAAATCCACTTGGCTTATCTGTGAGACTTGGGGCAGAGTTGATGCC 276
Qy 1130 TTCCCC--AGAGCCAGACCACTGGGGGTGATCATTTGGGGATTCTGCTAGAGTACTT 1186
Db 275 CTTCCTCCATGAGCCATGACCACTGGGGGTGATCATTTGGGGATTCTGCTAGAGTACTT 216
Qy 1187 TGATAGAGTGTGGGTGGGGGGGACTTGTGAGATGAGCTCACTTCTCCATCCC 1246
Db 215 TGATAGAGTGTGGGTGGGGGGGACTTGTGAGATGAGCTCACTTCTCCATCCC 156
Qy 1247 AGAAGCGGGGCTTAAAGCCACTTACAGTTTCACTCATGAAACCTTTGATCTTTGGT 1306
Db 155 AGAAGCGGGGCTTAAAGCCACTTACAGTTTCACTCATGAAACCTTTGATCTTTGGT 96
Qy 1307 GTCTGAGATTTATCTCT-GGGTGCTGCAATATCTGAGAGGAAGTAAAGAAATCAATC 1365
Db 95 GTCTGAGATTTATCTCTAGGGGTGCTGCAATATCTGAGAGGAAGTAAAGAAATCAATC 36
```

Qy 1366 TTGCTGCCCCCAGCTACACTACAGCTGGAGCC 1400  
Db 35 TTGCTGCCCCAACACTACAGCTTACAGCTGGAGCC 1

RESULT 10  
BE789709 805 bp mRNA linear EST 20-OCT-2000  
LOCUS 601481483F1 NIH\_MGC\_68 Homo sapiens cDNA clone IMAGE:3884156 5',  
DEFINITION mRNA sequence.  
ACCESSION BE789709  
VERSION BE789709.1 GI:10210907  
KEYWORDS EST.  
SOURCE human.  
ORGANISM Homo sapiens  
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.

REFERENCE 1 (bases 1 to 805)  
AUTHORS NIH-MGC <http://mgi.nci.nih.gov/>.  
TITLE National Institutes of Health, Mammalian Gene Collection (MGC)  
JOURNAL Unpublished (1999)  
COMMENT Contact: Robert Strausberg, Ph.D.  
Email: [cgabs-remail.nih.gov](mailto:cgabs-remail.nih.gov)  
Tissue Procurement: DCTD/DRP/Gazdar  
CDNA Library Preparation: Life Technologies, Inc.  
DNA Library Arrayed by: The I.M.A.G.E. Consortium (LNL)  
DNA sequencing by: Incyte Genomics, Inc.  
Clone distribution: MGC clone distribution information can be  
found through the I.M.A.G.E. Consortium/LNL at:  
<http://image.lnl.gov>  
Plate: LAMB657 row: e column: 21  
High quality sequence stop: 610.  
Location/Qualifiers

#### FEATURES

source  
1..805  
/organism="Homo sapiens"  
/db\_xref="taxon:9606"  
/clone="IMAGE:3884156"  
/clone\_id="NIH\_MGC\_68"  
/tissue\_type="large cell carcinoma"  
/lab\_host="DH10B (phage-resistant)"  
/note="Organ: lung; Vector: pCMV-SPORT6; Site 1: NotI,  
Site 2: SalI; Cloned unidirectionally. Primer: oligo dT.  
Average insert size 1.8 kb. Library constructed by Life  
Technologies."

BASE COUNT 134 a 236 c 291 g 144 t  
ORIGIN

Query Match 42.5%; Score 610; DB 12; Length 805;  
Best Local Similarity 91.0%; Pred. No. 2.9e-118;  
Matches 728; Conservative 0; Mismatches 60; Indels 12; Gaps 7;

```
Qy 356 TGGCGGCGCGCTGCGCCAGCACTGGTGGCGCTGCACTGAGCTGCGCGCGCGCG 415
Db 16 TGGAGCTGGGCAAAATAGCACTGGTGGCGCTGCACTGAGCTGCGCGCGCGCG 75
Qy 416 AGCGGCTGAGCGCTTGTGTGGCGAGAGAGCGCTGTTTGAATTGATCTTACGCCAGC 475
Db 76 AGCGGCTGAGCGCTTGTGTGGCGAGAGAGCGCTGTTTGAATTGATCTTACGCCAGC 135
Qy 476 AGCCGAGCCGGCTCCGGGATGAGAACTGGCTGAGCTGAGAGAGAGTCCGTGGGAAATCGA 535
Db 136 AGCCGAGCCGGCTCCGGGATGAGAACTGGCTGAGCTGAGAGAGAGTCCGTGGGAAATCGA 195
Qy 536 AGTGGGCTGGGGGCGCGGGGTGGCGAGCGGGAGGTGCTTTCGCCCCCTTGCAGCCCC 595
Db 136 AGTGGGCTGGGGGCGCGGGGTGGCGAGCGGGAGGTGCTTTCGCCCCCTTGCAGCCCC 255
Qy 596 CGGTGCTTCTGTGTGAGAGTGAAGCGCGCGCGCGCCACTGCTCCAGACTTTTC 655
Db 256 CGGTGCTTCTGTGTGAGAGTGAAG--GCCGCGCGCGCGCGCACTCCAGACTTTTC 312
Qy 656 TGTTCAGAGGTACGCTGTAGTAATCGGCGCGCTGAGCTGAGAGCAACAGAGCTTGG 715
```

```
Db 313 TGTTCACACCGCTGTAGTATGATCGCGCTGAGGCTGTAAAGACCAACAGACGTTG 371
Qy 716 CGCGCTCTGTGGGCTCTAAATAGGCGCAAGTGGGGCGCTCACTGACGAGGAGCTGGCGG 775
Db 372 CGCGCTCTGTGGGCTCTAAATAGGCGCAAGTGGGGCGCTCACTGACGAGGAGCTGGCGG 431
Qy 776 CGCTGCGGAGCCCGGCGCTGAGCTGCTGAGCTTACAGATACAGAGCGGAGGAGCTTACG 835
Db 432 CGCTGCGGAGCCCGGCGCTGAGCTGCTGAGCTTACAGATACAGAGCGGAGGAGCTTACG 490
Qy 836 AGCAGGCTTTCACAGTGTGTGGGGCTTCTGTCAGGCGGAGGCGCGCCACAGCTGC 895
Db 491 AGCAGGCTTTCACAGTGTGTGGGGCTTCTGTCAGGCGGAGGCGCGCCACAGCTGC 550
Qy 896 AGCGCTGTGTGAGGAGCACTGAGGAGAAAGAGCTCACAGCTGAGGAGGAGGAGCTTGTG 955
Db 551 AGCGCTGTGTGAGGAGCACTGAGGAGAAAGAGCTCACAG-CTGAGCAGAGGAGCTTGTG 609
Qy 956 GCGTACCGATCCCAATAGCGCGCTGAGCAGGAGGAGTGCAGCCAGCTTTTGAGAA 1015
Db 610 GCGTACCGATCCCAATAGCGCGCTGAGCAGGAGGAGTGCAGCCAGCTTTTGAGAA 665
Qy 1016 CCGTGAATGCTTGAAGGTTCTTCTGCGGCTATGCTGAGACCCCTGTCATTCAGGAG 1075
Db 666 CCGTGAATGCTTGAAGGTTCTTCTGCGGCTATGCTGAGACCCCTGTCATTCAGGAG 724
Qy 1076 CCGTGAATGCTTGAAGGTTCTTCTGCGGCTATGCTGAGACCCCTGTCATTCAGGAG 1134
Db 725 CCGTGAATGCTTGAAGGTTCTTCTGCGGCTATGCTGAGACCCCTGTCATTCAGGAG 784
Qy 1135 CAGGAGCCGACGACCTGAGG 1154
Db 785 AAGGAGCCGACGACCTGAGG 804
```

```
RESULT 11
LOCUS BG717460 632 bp mRNA linear EST 08-MAY-2001
DEFINITION 602689219F1 NIH_MGC_97 Homo sapiens cDNA clone IMAGE:4821509 5',
mRNA sequence.
ACCESSION BG717460
VERSION BG717460.1 GI:13996647
KEYWORDS EST.
SOURCE human.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
REFERENCE 1 (bases 1 to 632)
AUTHORS NIH-MGC http://mgc.nci.nih.gov/.
TITLE National Institutes of Health, Mammalian Gene Collection (MGC)
JOURNAL Unpublished (1999)
COMMENT Contact: Robert Straubeberg, Ph.D.
Email: rstraub@mail.nih.gov
Tissue Procurement: Miklos Palkovics, M.D., Ph.D.
cDNA Library Preparation: Michael J. Brownstein (NHGRI), Shiroki
Toshiyuki and Piero Carninci (RIKEN)
CDNA Library Arrayed by: The I.M.A.G.E. Consortium (LNL)
DNA Sequencing by: Incyte Genomics, Inc.
Clone distribution: MGC clone distribution information can be
found through the I.M.A.G.E. Consortium/LNL at:
http://image.llnl.gov
Plate: LLM0728 row: f column: 06
High quality sequence stop: 632.
Location/Qualifiers
```

```
FEATURES
Source
1. 632
/organism="Homo sapiens"
/db xref="taxon:9606"
/clone="IMAGE:4821509"
/clone_1db="NIH_MGC_97"
/lab_host="DH10B"
/note="Organ: testis; Vector: pBluescript (modified
pBluescript KS+); Site_1: BamHI; Site_2: SalI-XhoI (gtcgag
```

```
); Oligo-dT primed using primer 5'-TTTTTTTTTTTTTNN-3',
size-selected for average insert size 2.2 kb and
normalized to 10^5. This is a primary library enriched
for full-length clones and constructed using the
cap-trapper method (Carninci, in preparation). Library
constructed by M. Brownstein (NHGRI, National
Institutes of Health). Note: this is a NIH_MGC Library."
BASE COUNT 97 a 195 c 249 g 91 t
ORIGIN
```

```
Query Match 42.2%; Score 605; DB 12; Length 632;
Best Local Similarity 100.0%; Pred. No. 3.2e-117;
Matches 605; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 CTGGCGGCGCGTGGGAACTCCAGGCGCGCGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 60
Db 28 CTGGCGGCGCGTGGGAACTCCAGGCGCGCGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 87
Qy 61 CAAATATGGGACGAGAGAGTGGTGGGAGCGCATACCTGTTTGTGAGTCTGCTTGAGAC 120
Db 88 CAAATATGGGACGAGAGAGTGGTGGGAGCGCATACCTGTTTGTGAGTCTGCTTGAGAC 147
Qy 121 AAGTGTCTCTGTGGATGCTTACGCGGACCCCGAGCAGAGGAGGAGGAGGAGGAGGAGGAGG 180
Db 148 AAGTGTCTCTGTGGATGCTTACGCGGACCCCGAGCAGAGGAGGAGGAGGAGGAGGAGG 207
Qy 181 CTGAGGCTGCTTGGGAGAGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 240
Db 208 CTGAGGCTGCTTGGGAGAGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 267
Qy 241 CACCGCAGCAGCCCGCAGCTGATGTCAGTCTGATTCGCGGCGGAGGAGGAGGAGGAGGAGG 300
Db 268 CACCGCAGCAGCCCGCAGCTGATGTCAGTCTGATTCGCGGCGGAGGAGGAGGAGGAGGAGG 327
Qy 301 CGCTTCTCTCGCGCTTACCGCGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 360
Db 328 CGCTTCTCTCGCGCTTACCGCGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 387
Qy 361 GCGCGCTGCGCCAGCAGCTGAGTGGCGCTGCACTGAGGAGTGGCGCGCGCGCGCGCGAGG 420
Db 388 GCGCGCTGCGCCAGCAGCTGAGTGGCGCTGCACTGAGGAGTGGCGCGCGCGCGCGCGAGG 447
Qy 421 CTGAGCCTTGTGCTGCGGAGCAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 480
Db 448 CTGAGCCTTGTGCTGCGGAGCAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 507
Qy 481 GACCGGCTCCGGAGTGAAGACTGCTGAGCTGAGGAGTGGCTGCGAATCTGAAGTGC 540
Db 508 GACCGGCTCCGGAGTGAAGACTGCTGAGCTGAGGAGTGGCTGCGAATCTGAAGTGC 567
Qy 541 GCGTGGGCGGCGCGGAGTGGCGAGCGGAGGAGTGGCTTGGCGCCCTTGGAGCGCCCGGAG 600
Db 568 GCGTGGGCGGCGCGGAGTGGCGAGCGGAGGAGTGGCTTGGCGCCCTTGGAGCGCCCGGAG 627
Qy 601 CCTTC 605
Db 628 CCTTC 632

RESULT 12
LOCUS A1144401 609 bp mRNA linear EST 26-OCT-1998
DEFINITION db77f02.x1 Soares fetal heart MbH19w Homo sapiens cDNA clone
IMAGE:1706139.3 similar to SWTRAD HUMAN Q15628 TUMOR NECROSIS
FACTOR RECEPTOR TYPE 1 ASSOCIATED PROTEIN ; mRNA sequence.
ACCESSION A1144401
VERSION A1144401.1 GI:3666210
KEYWORDS EST.
SOURCE human.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
REFERENCE 1 (bases 1 to 609)
```

**AUTHORS** NCI-CGAP <http://www.ncbi.nlm.nih.gov/ncicgap>.  
**TITLE** National Cancer Institute, Cancer Genome Anatomy Project (CGAP),  
**JOURNAL** Tumor Gene Index  
**COMMENT** Unpublished (1997)  
 Contact: Robert Strausberg, Ph.D.  
 Email: [cgapbs-ri@mail.nih.gov](mailto:cgapbs-ri@mail.nih.gov)  
 This clone is available royalty-free through LNL; contact the  
 IMAGE Consortium ([info@image.llnl.gov](mailto:info@image.llnl.gov)) for further information.  
 Insert Length: 1137 Std Error: 0.00  
 Seq primer: -40m13 fwd. ET from Amerham  
 High quality sequence stop: 463.  
**FEATURES**  
 source  
 1..609  
 /organism="Homo sapiens"  
 /db\_xref="taxon:9606"  
 /clone="IMAGE:1706139"  
 /clone\_lib="Soares\_fetal\_heart\_NbH19W"  
 /sex="unknown"  
 /dev\_stage="19 weeks"  
 /lab\_host="DH10B (ampicillin resistant)"  
 /note="Organ: heart; Vector: pT73D (Pharmacia) with a  
 modified polylinker; Site 1: Not I; Site 2: Eco RI; 1st  
 strand cDNA was primed with a Not I - oligo(dT) primer [5'  
 TGTACCAATCTGAAGTGGAGCGCGCCGATCTTTTCTTTTCTTTT 3'],  
 double-stranded cDNA was size selected, ligated to Eco RI  
 adapters (Pharmacia), digested with Not I and cloned into  
 the Not I and Eco RI sites of a modified pT73 vector  
 (Pharmacia). Library went through one round of  
 normalization to a Cot = 5. Library constructed by  
 M. Fatima Bonaldo. This library was constructed from the  
 same fetus as the fetal lung library, Soares fetal lung  
 NbH19W."  
**BASE COUNT** 90 a 198 c 224 g 97 t  
**ORIGIN**  
 Query Match 42.0%; Score 602.6; DB 9; Length 609;  
 Best Local Similarity 99.3%; Pred. No. 1e-116;  
 Matches 605; Conservative 0; Mismatches 4; Indels 0; Gaps 0;  
 Oy 376 CACTGGTGGCCGCTGCACTGAGAGCGCGCGCGCGCGCGCTTGACCTTTGCTG 435  
 Db 1 CACTGGTGGCCGCTGCACTGAGAGCGCGCGCGCGCGCGCTTGACCTTTGCTG 60  
 Oy 436 GCGGACGAGAGCGCTGTTGATTGATTCATCTGACCCAGCCGCGCTCCGGGAT 495  
 Db 61 GCGGACGAGAGCGCTGTTGATTGATTCATCTGACCCAGCCGCGCTCCGGGAT 120  
 Oy 496 GAGGAACTGGCTGAGCTGAGAGATGCGCTGCCAATCTGAAGTGGCGCTCGGGGCGCCG 555  
 Db 121 GAGGAACTGGCTGAGCTGAGAGATGCGCTGCCAATCTGAAGTGGCGCTCGGGGCGCCG 180  
 Oy 556 GGTGGCGAGCGGAGAGTTCCTTCGCGCCCTTGACGCGCGCGCGCGCTTCGTGCGAG 615  
 Db 181 GGTGGCGAGCGGAGAGTTCCTTCGCGCCCTTGACGCGCGCGCGCGCTTCGTGCGAG 240  
 Oy 616 GTGAAGCCCGCCCGCGCGCGCGCGCGCGCGCGCGCGCGCGCTTCGTGCGAGGATGAGCTGTA 675  
 Db 241 GTGAAGCCCGCCCGCGCGCGCGCGCGCGCGCGCGCGCGCGCTTCGTGCGAGGATGAGCTGTA 300  
 Oy 676 GTGAATCGGCGCGCTGAGCTGAAAGCAACAGACGTTGCGCGCTCTGTGGGTCTCAAA 735  
 Db 301 GTGAATCGGCGCGCTGAGCTGAAAGCAACAGACGTTGCGCGCTCTGTGGGTCTCAAA 360  
 Oy 736 TGGCCGCAAGTGGGGGCGCTCATCTGCGAGCGAGGCTGCGGGCGCGTGGGGAGACCGGGGCTG 795  
 Db 361 TGGCCGCAAGTGGGGGCGCTCATCTGCGAGCGAGGCTGCGGGCGCGTGGGGAGACCGGGGCTG 420  
 Oy 796 GACTGGCTGGCTTCAAGTACGAGCGCGAGGAGCTGTACGAGAGCGCTTCCAGCTGCTG 855  
 Db 421 GACTGGCTGGCTTCAAGTACGAGCGCGAGGAGCTGTACGAGAGCGCTTCCAGCTGCTG 480  
 Oy 856 CGGCGCTTGTGCGAGGCGAGAGGCGCGCGCGCGCGCGCTGCGAGCGCTGTGTGAGGAGCACTC 915

Db 481 CGGCGCTTGTGCGAGGCGAGGCGCGCGCGCGCGCGCGCTGCGAGCGCTGTGTGAGGAGCACTC 540  
 Oy 916 GAGGAGAACGAGCTCACACGAGCTGGAGAGAGCTTGGGCTGACCGATCCCAATGGC 975  
 Db 541 GAGGAGAACGAGCTCACACGAGCTGGAGAGAGCTTGGGCTGACCGATCCCAATGGC 600  
 Oy 976 GGCTTGGCC 984  
 Db 601 GGCTTGGCC 609  
**RESULT 13**  
 A1439047 609 bp mRNA linear EST 28-MAR-1999  
 LOCUS c89h08.x1 NCI CGAP CLL1 Homo sapiens cDNA clone IMAGE:2073375 3'  
 DEFINITION similar to SW:TMAD HUMAN 015628 TUMOR NECROSIS FACTOR RECEPTOR TYPE  
 1 ASSOCIATED PROTEIN; mRNA sequence.  
 ACCESSION A1439047  
 VERSION A1439047.1 GI:4301586  
 KEYWORDS EST.  
 SOURCE human.  
 ORGANISM Homo sapiens  
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.  
**REFERENCE**  
 1 (bases 1 to 609)  
**AUTHORS** NCI-CGAP <http://www.ncbi.nlm.nih.gov/ncicgap>.  
**TITLE** National Cancer Institute, Cancer Genome Anatomy Project (CGAP),  
**JOURNAL** Tumor Gene Index  
**COMMENT** Unpublished (1997)  
 Contact: Robert Strausberg, Ph.D.  
 Email: [cgapbs-ri@mail.nih.gov](mailto:cgapbs-ri@mail.nih.gov)  
 Tissue Procurement: Ash Alizadeh, John Byrd, M.D., Mike Grever,  
 M.D., Louis M. Staudt, M.D., Ph.D.  
 cDNA Library Preparation: M. Bento Soares, Ph.D.  
 cDNA Library Arrayed by: Greg Lennon, Ph.D.  
 DNA Sequencing by: Washington University Genome Sequencing Center  
 Clone distribution: NCI-CGAP clone distribution information can be  
 found through the I.M.A.G.E. Consortium/LNL at:  
 www-bio.llnl.gov/bdrp/image/image.html  
 Insert Length: 1151 Std Error: 0.00  
 Seq primer: -40UP from Gibco  
 High quality sequence stop: 416.  
**FEATURES**  
 source  
 1..609  
 /organism="Homo sapiens"  
 /db\_xref="taxon:9606"  
 /clone="IMAGE:2073375"  
 /clone\_lib="NCI-CGAP-CLL1"  
 /tissue\_type="B-cell, chronic lymphocytic leukemia"  
 /lab\_host="DH10B"  
 /note="Vector: pT73D-Pac (Pharmacia) with a modified  
 polylinker; Site 1: Not I; Site 2: Eco RI; 1st strand cDNA  
 was primed with a Not I - oligo(dT) primer [5'  
 TGTACCAATCTGAAGTGGAGCGCGCGCGCGCGCGCTTTTCTTTTCTTTT  
 T 3']; double-stranded cDNA was ligated to Eco RI  
 adapters (Pharmacia), digested with Not I and cloned into  
 the Not I and Eco RI sites of the modified pT73 vector.  
 Library is normalized, and was constructed by Bento  
 Soares and M. Fatima Bonaldo."  
**BASE COUNT** 91 a 200 c 222 g 96 t  
**ORIGIN**  
 Query Match 41.9%; Score 601; DB 9; Length 609;  
 Best Local Similarity 99.2%; Pred. No. 2.2e-116;  
 Matches 604; Conservative 0; Mismatches 5; Indels 0; Gaps 0;  
 Oy 376 CACTGGTGGCCGCTGCACTGAGAGCGCGCGCGCGCGCGCTTGACCTTTGCTG 435  
 Db 1 CACTGGTGGCCGCTGCACTGAGAGCGCGCGCGCGCGCGCTTGACCTTTGCTG 60  
 Oy 436 GCGGACGAGAGCGCTGTTGATTGATTCATCTGACCCAGCGCGCGCTCCGGGAT 495  
 Db 61 GCGGACGAGAGCGCTGTTGATTGATTCATCTGACCCAGCGCGCGCTCCGGGAT 120

QY	496	GAAGAATCGGCTGACCTGGAGAGATGCGCTGCGAAATCTGAAATGTGGGGCTCGGGGGCCGG	555
Db	121	GAAGAATCGGCTGACCTGGAGAGATGCGCTGCGAAATCTGAAATGTGGGGCTCGGGGGCCGG	180
QY	556	GCTGGCGACGCGGGAGGTGCTTGTGGCCCTTTCAGAGCCCGGTCGCTCTCTGTGAG	615
Db	181	GCTGGCGACGCGGGAGGTGCTTGTGGCCCTTTCAGAGCCCGGTCGCTCTCTGTGAG	240
QY	616	GTGAAGCCGCGCGCCCGCCGCTGACCTTGTCCAGACTTTTCTGTCCAGGCTAGCCTGT	675
Db	241	GTGAAGCCGCGCGCCCGCCGCTGACCTTGTCTTGTCCAGGCTAGCCTGT	300
QY	676	GTGAATCGGCGCGCTGAGCCTGAAAGACCAAGAGCTTGGCGGCTGTGGGTCTCAA	735
Db	301	GTGAATCGGCGCGCTGAGCCTGAAAGACCAAGAGCTTGGCGGCTGTGGGTCTCAA	360
QY	736	TGGCCCAAGGTGGGGCGCTCACTGTGACGAGGCTCCGCGCGCTGCGGACCCGCGCTG	795
Db	361	TGGCCCAAGGTGGGGCGCTCACTGTGACGAGGCTCCGCGCGCTGCGGACCCGCGCTG	420
QY	796	GACTGCTGCGCTTACGATGACGAGGCGGACGAGGACTGTACAGACAGCCTTCCAGCTGCTG	855
Db	421	GACTGCTGCGCTTACGATGACGAGGCGGACGAGGACTGTACAGACAGCCTTCCAGCTGCTG	480
QY	856	CGGCGCTTGTGTACAGGCGGAGGCGCGCGCGCGCAAGCTGACGCGCTGTGGAGGCACTC	915
Db	481	CGGCGCTTGTGTACAGGCGGAGGCGCGCGCGCGCAAGCTGACGCGCTGTGGAGGCACTC	540
QY	916	GAGGAGAACGAGCTCACCAGCCTGTGACAGAGACTTGTGGGCTGACCGATCCAAATGAC	975
Db	541	GAGGAGAACGAGCTCACCAGCCTGTGACAGAGACTTGTGGGCTGACCGATCCAAATGAC	600
QY	976	GAGGCTGGCC 984	
Db	601	GAGGCTGGCC 609	
RESULT 14			
LOCUS	AK018592	1688 bp	linear
DEFINITION	Mus musculus adult male cecum cDNA, RIKEN full-length enriched library, clone:9130005N23;homolog to TUMOR NECROSIS FACTOR RECEPTOR TYPE 1 ASSOCIATED DEATH DOMAIN PROTEIN (TNFR1-ASSOCIATED DEATH DOMAIN PROTEIN), full insert sequence.		
ACCESSION	AK018592	GI:12858377	
VERSION	AK018592.1		
KEYWORDS	HTC; CAP trapper.		
SOURCE	Mus musculus (strain:C57BL/6J) adult male cecum cDNA to mRNA, clone:11b;RIKEN full-length enriched mouse cDNA library clone:9130005N23.		
ORGANISM	Mus musculus		
REFERENCE	Bukaryota; Metazoa; Chordata; Craniata; Vertebrate; Euteleostomi; Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.		
AUTHORS	Carninci, P. and Hayashizaki, Y.		
TITLE	High-efficiency full-length cDNA cloning		
JOURNAL	Meth. Enzymol. 303, 19-44 (1999)		
PMID	99279253		
PMID	10349636		
REFERENCE			
AUTHORS	Carninci, P., Shibata, Y., Hayatsu, N., Sugahara, Y., Shibata, K., Itoh, M., Kono, H., Okazaki, Y., Muramatsu, M. and Hayashizaki, Y.		
TITLE	Normalization and subtraction of cap-trapper-selected cDNAs to prepare full-length cDNA libraries for rapid discovery of new genes		
JOURNAL	Genome Res. 10 (10), 1617-1630 (2000)		
PMID	10499374		
PMID	11042159		
REFERENCE			
AUTHORS	Shibata, K., Itoh, M., Aizawa, K., Nagaoka, S., Sasaki, N., Carninci, P., Kono, H., Akiyama, J., Nishi, K., Kitsuina, T., Tashiro, H., Itoh, M., Sumi, N., Ishii, Y., Nakamura, S., Hazama, M., Nishino, T., Harada, A., Yamamoto, K., Matsumoto, H., Sakaguchi, S., Ikegami, T., Kaeshiwa, K.,		

FEATURES

TITLE

JOURNAL

PUBLISHED

REFERENCE

AUTHORS

4 Kawai, J., Shingawa, A., Shibata, K., Yoshino, M., Itoh, M., Ishii, Y., Arakawa, T., Hara, A., Fukunishi, Y., Konno, H., Aadahi, J., Fukuda, S., Aizawa, K., Izawa, M., Nishi, K., Kiyosawa, H., Kondo, S., Yamataka, I., Saito, T., Okazaki, Y., Gojobori, T., Bono, H., Kasukawa, T., Saito, R., Kadota, K., Matsuda, H., Ashburner, M., Batalov, S., Casavant, T., Fleischmann, W., Gaasterland, T., Gissi, C., King, B., Kochwa, H., Kiehl, P., Lewis, S., Matsuo, Y., Nikaido, I., Pesole, G., Quackenbush, J., Schiraldi, L.M., Staab, J., Suzuki, R., Tomita, M., Wagner, L., Mashio, T., Sakai, K., Okido, T., Furuno, M., Aono, H., Baldarelli, R., Barsh, G., Blake, J., Boffelli, D., Bojunga, N., Carninci, P., de Bona, M., Brownstein, M., J., Bult, C., Fletcher, C., Fujita, M., Gariboldi, M., Gustincich, S., Hill, D., Hofmann, M., Hume, D., Kamaya, M., Lee, N., Lyons, P., Marchionni, L., Mashima, J., Mazzarelli, D., Mombauer, P., Nordone, P., Rung, B., Ringwald, M., Rodriguez, I., Sakamoto, N., Sasaki, H., Sato, K., Schonbach, C., Seya, T., Shibata, Y., Storch, K.F., Suzuki, H., Toyono-Oka, K., Wang, K.H., Welt, K., Whitaker, C., Wilming, L., Wysshaw-Boris, A., Yoshida, K., Hasegawa, Y., Kawai, J., Kohetsuki, S. and Hayashizaki, Y.

Functional annotation of a full-length mouse cDNA collection

Nature 409 (6821), 685-690 (2001)

21085660

11217851

5 (bases 1 to 1688)

Aadahi, J., Aizawa, K., Akahira, S., Akimura, T., Aono, H., Arai, A., Arakawa, T., Baldarelli, R., Bono, H., Brownstein, M., Bult, C., Carninci, P., Fukuda, S., Fukunishi, Y., Furuno, M., Hanagaki, T., Hara, A., Hayatsu, N., Hill, D., Hiramoto, K., Hiraoka, T., Hori, F., Hume, D., Imotani, K., Ishii, Y., Itoh, M., Izawa, M., Kasukawa, T., Kato, H., Kawai, J., Kojima, Y., Konno, H., Kouda, M., Koya, S., Kunitara, C., Matsuyama, T., Miyazaki, A., Nishi, K., Nomura, K., Numata, R., Ono, M., Okazaki, Y., Okido, T., Owa, C., Quackenbush, J., Saito, H., Saito, R., Sakai, C., Sakai, K., Sano, H., Saeki, D., Schiraldi, L., Shibata, K., Shibata, Y., Shingawa, A., Shiraki, T., Sogabe, Y., Suzuki, H., Tagami, M., Tagawa, A., Takahashi, F., Tanaka, T., Teijima, Y., Toyota, T., Yamamura, T., Yamataka, I., Yasunishi, A., Yoshida, K., Yoshino, M., Muramatsu, M. and Hayashizaki, Y.

Direct Submission

Submitted (10-JUL-2000) Yoshinori Hayashizaki, The Institute of Physical and Chemical Research (RIKEN), Laboratory for Genome Exploration and Research Group, RIKEN Genomic Sciences Center (GSC), RIKEN Yokohama Institute; 1-7-22 Suehiro-cho, Tsurumi-ku, Yokohama, Kanagawa 230-0045, Japan (E-mail: genome-rsg@gsc.riken.go.jp, URL: <http://genome.gsc.riken.go.jp/>, Tel: 81-45-503-9222, Fax: 81-45-503-9216)

PLEASE VISIT OUR WEB SITE (<http://genome.gsc.riken.go.jp/>) FOR FURTHER DETAILS.

cDNA library was prepared and sequenced in Mouse Genome Encyclopaedia Project of Genome Exploration Research Group in Riken Genomic Sciences Center and Genome Science Laboratory in RIKEN. Division of Experimental Animal Research in Riken contributed to prepare mouse tissues. First strand cDNA was primed with a primer [5' GAGAGAGAGATCCGACGATTAATTAATTAATTCGCCGCCGCC 3']. cDNA was prepared by using triazole thermo-activated reverse transcriptase and subsequently enriched for full-length by cap-trapper. cDNA went through one round of normalization to Rot = 10.0 and subtraction to Rot = 185.2. Second strand cDNA was prepared with the primer adapter of sequence [5' GAGAGAGATTCCTCGATTAATTAATTAATTCGCCGCCGCC 3']. cDNA was cleaved with BamHI and XhoI. Vector: a modified plasmid vector KS(+) after bulk excision from lambda FLC I. Cloning sites, 5' end: SalI, 3' end: BamHI. Host: DH10B.

Location/Qualifiers





```
Db      1  GGAGCCCTGTGCGCGCTTCTCCGCGCTACCGCGAGGGGGCGCTGCGCGCGCTGCG
QY      347  AGAGAGCCTTG3CG3GCGCGCTCGCCAGACACTGCGTGCGCTGCAACTGAGCTGCGG 406
      |||
Db      61  AGAGAGCCTTGCGCGCGCGCTCGCCAGACACTGCGTGCGCTGCAACTGAGCTGCGG 120
      |||
QY      407  CCGGCGCGCGAGCGGC-TGACGCTTTGCTG3GCGGACGAGAGCGCTGTTGAGTTGCATC 465
      |||
Db      121  CCGGCGCGCGAGCGCGCTGACGCTTTGCTGCGCGGACGAGGAGCGCTGTTGAGTTGCATC 180
      |||
QY      466  CTAGCCCGAGAGCGCCGACCGGCTCCGGGATGAAAGACTGCGTGAAGATGCGCTG 525
      |||
Db      181  CTAGCCCGAGAGCGCCGACCGGCTCCGGGATGAAAGACTGCGTGAAGATGCGCTG 240
      |||
QY      526  CGAAATCTGAAGTGC3GCGCTCG3GG3GCGCGAGCGAGAGTGCCTTCG3GCGCGCC 585
      |||
Db      241  CGAAATCTGAAGTGC3GCGCTCG3GG3GCGCGAGCGAGAGTGCCTTCG3GCGCGCC 300
      |||
QY      586  TTGCGCGCGCGCGGCTCTGTGCGAGGTGAAGCGCGCGCGCGCGCGCACTGCG 645
      |||
Db      301  TTGCGCGCGCGCGGCTCTGTGCGAGGTGAAGCGCGCGCGCGCGCGCA-CTGCC 359
      |||
QY      646  CAGACTTTTCTGTTCCAGGGGTCAAGCTGTGTA-GTGAATGCGCGCGCTGAAGCACTG 704
      |||
Db      360  CAGACTTTGCTGTACAGGGGTCAAGCTGTGTA-GTGAATGCGCGCGCTGAAGCACTG 419
      |||
QY      705  ACAGACGTTGCGCGCGCT-CTGTGGGTCTCAAAATGCGCAAGTGG3GCGCTCACTGAGC 763
      |||
Db      420  ACAGACGTTGCGCGCGCTCTGTGGGTCTCAAAATGCGCAAGTGG3GCGCTCACTGAGC 479
      |||
QY      764  GAGGCTGCGCGCGCGCTGCGGACCGCGCGCTGAGCT--CGTGGCTTACGATGAGCG 821
      |||
Db      480  GAGGCTGCGCGCGCGCTGCGGACCGCGCGCTGAGCTTCCGTTGGCTTACGATGAGCG 539
      |||
QY      822  C-GAGGACTGTAGAGACAGGCTTTCAG-CTGCTGCGCGCGCTTGTGCAAGCGCGAGGCG 879
      |||
Db      540  CGAGGGGACTGTAGAGACAGGCTTTCAGCTGTGCGCGCGCTTGTGCAAGCGCGAGGCG 599
      |||
QY      880  CGCGCGCGCAAGCTGCGAGCGCTGTGGAGGCACTCGAGG-AGAAAGAGCTCAACAGCT 938
      |||
Db      600  CGCGCGCGCGCAAGCTGAGCGCGCTGTGGAGGCACTCGAGGAGAGAGCTCAACAGCT 659
      |||
QY      939  -GGCAGAGGACTTGTGCGCGCTGAGC--GATCCCAATGCGCGCGCTG3-CTAGACCA-- 991
      |||
Db      660  GGGGAGAGGACTGTTGGGCTTTCAGCGATTCCCAATGCGCGCGCTG3GCTAGACCAAG 719
      |||
QY      992  -GGGCTGCAAGCAAGCTTTTGAAGAACTGATGCGCTTGAAGGTTCTTGTGCGGCTATTG 1050
      |||
Db      720  TGGGTTGCAAGCAAGTTTGGGAAAACTGATGCGCTTCAGG3GTTCTTGTGCGGATATG 779
      |||
QY      1051  CTGAACCCCTGTGCATGCAAGGAGCCCTGAACCTGCACTTG3GCTTATCTGTGAGCCTGC 1110
      |||
Db      780  GCTGAACCTGTGCTCATCAC-GGACCTGAACCTC--AGTGGCTTATCTGTGAGC--TG 834
      |||
QY      1111  TGGGCGAGAGTTGCTTTCCTCCAGAGGCA 1143
      |||
Db      835  CTGGCGAAAGTGAATGCTTCCCGGAGGCA 867
      |||
```

Search completed: February 3, 2003, 16:12:57  
Job time : 2225 secs

**THIS PAGE BLANK (USPTO)**



361 GCCGCGCTCGCCAGCACTCGGTGCGCTGCAACTGAGCTGCGCGCGCCAGCGG 420  
QY CTGACGCGCTTTGCTGCGGAGAGAGCGCTGTTGAGTTGATCTTGACCCAGAGCC 480  
Db CTGACGCGCTTTGCTGCGGAGAGAGCGCTTGTGAGTTGATCTTGACCCAGAGCC 480  
QY GACCGCGCTCGGAGATGAAGAACTGCTGAGCTGAGAGATGCGCTGGAATCTGAATGTC 540  
Db GACCGCGCTCGGAGATGAAGAACTGCTGAGCTGAGAGATGCGCTGGAATCTGAATGTC 540  
QY GAGCTGCGGCGCGCGGAGCGAGCGGAGGTCGCTTCCGCCCCCTTGACAGCCCCCGGTG 600  
Db GAGCTGCGGCGCGCGGAGCGAGCGGAGGTCGCTTCCGCCCCCTTGACAGCCCCCGGTG 600  
QY CCCTCTGCTGCGGAGGTGAAGCGCGCGCGCGCCAGCTGCCAGCTTTCTGTTTC 660  
Db CCCTCTGCTGCGGAGGTGAAGCGCGCGCGCGCCAGCTGCCAGCTTTCTGTTTC 660  
QY CAGGCTCAGCGCTGATGATGAACTGCGCGCTGAGCTTGAAGAGCAACAGAGCTTCCGCGC 720  
Db CAGGCTCAGCGCTGATGATGAACTGCGCGCTGAGCTTGAAGAGCAACAGAGCTTCCGCGC 720  
QY TCTGTGGGTCTCAAAATGCGCGAGGTCGCGCTCACTGACGAGGCTGCCGCGCGCTG 780  
Db TCTGTGGGTCTCAAAATGCGCGAGGTCGCGCTCACTGACGAGGCTGCCGCGCGCTG 780  
QY CGGAGCCCGGCGCTGAGACTCGCTGCGCTAACAGTACGAGCGCGAGAGGACTGTACAGAGC 840  
Db CGGAGCCCGGCGCTGAGACTCGCTGCGCTAACAGTACGAGCGCGAGAGGACTGTACAGAGC 840  
QY GCGCTTCCAGCTGCTGCGCGCGCTTCCGTCAGAGCGCGCGCGCGCGCGCTGACGCGC 900  
Db GCGCTTCCAGCTGCTGCGCGCGCTTCCGTCAGAGCGCGCGCGCGCGCGCTGACGCGC 900  
QY CTGTGTGAGGCACTGAGAGAGCAAGAGCTCACAGCTTGGCAGAGACTTGTGCGCGCTG 960  
Db CTGTGTGAGGCACTGAGAGAGCAAGAGCTCACAGCTTGGCAGAGACTTGTGCGCGCTG 960  
QY ACCGATCCCAATGAGGCGCTGCGCTGAGACAGAGGTCGAGCGAGCTTTTGGAGAACCTG 1020  
Db ACCGATCCCAATGAGGCGCTGCGCTGAGACAGAGGTCGAGCGAGCTTTTGGAGAACCTG 1020  
QY ATGCGCTTAGGGTCTCTTCTGCGCGCTAATGCTGAACCCCTGTCACACAGAGCCCTGA 1080  
Db ATGCGCTTAGGGTCTCTTCTGCGCGCTAATGCTGAACCCCTGTCACACAGAGCCCTGA 1080  
QY AACTCCACTTGGCTATGCTGCTGAGACTGCTGCGGCAAGATTGATGCTTCCAGAGAG 1140  
Db AACTCCACTTGGCTATGCTGCTGAGACTGCTGCGGCAAGATTGATGCTTCCAGAGAG 1140  
QY CCAGACCACTGGGCGGTGATCAATTGGGGAATTGCTCAGAGTACTTTGATATAGTGTGG 1200  
Db CCAGACCACTGGGCGGTGATCAATTGGGGAATTGCTCAGAGTACTTTGATATAGTGTGG 1200  
QY GTGGGGGGGAGACTTGTGAGAGATCAGCTCAGCTTCTCCCAACCCAGAGGCGGGCTTA 1260  
Db GTGGGGGGGAGACTTGTGAGAGATCAGCTCAGCTTCTCCCAACCCAGAGGCGGGCTTA 1260  
QY CAGCAGCCCTTACAGTTTCACTCATGAAGCACTTGATCTTTGTGTCTGAGACTTCAT 1320  
Db CAGCAGCCCTTACAGTTTCACTCATGAAGCACTTGATCTTTGTGTCTGAGACTTCAT 1320  
QY CCGTGGTGTCTGAGATCTGAGAGTGAAGTAAACAGAGATCAATCTTGTGCTGCGCCG 1380  
Db CCGTGGTGTCTGAGATCTGAGAGTGAAGTAAACAGAGATCAATCTTGTGCTGCGCCG 1380  
QY TCACACTCAGCGTGGAGCCCGAGATGTTAAAGCATGATTAATAAGTAAACAGCG 1435  
Db TCACACTCAGCGTGGAGCCCGAGATGTTAAAGCATGATTAATAAGTAAACAGCG 1435

RESULT 2  
US-08-414-625-1

Sequence 1, Application US/08414625  
Patent No. 5563039  
GENERAL INFORMATION:  
APPLICANT: Goedel, David V.  
APPLICANT: Hsu, Hailing  
TITLE OF INVENTION: INTRACELLULAR SIGNALING PROTEINS AND  
METHODS OF USE  
NUMBER OF SEQUENCES: 5  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: FLEHR, ROEBACH, TEST, ALBRITTON & HERBERT  
STREET: 4 Embarcadero Center, Suite 3400  
CITY: San Francisco  
STATE: California  
COUNTRY: USA  
ZIP: 94111-4187  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: PatentIn Release #1.0, Version #1.30  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/414,625  
FILING DATE:  
CLASSIFICATION: 435  
ATTORNEY/AGENT INFORMATION:  
NAME: Osman, Richard A  
REGISTRATION NUMBER: 36,627  
REFERENCE/DOCKET NUMBER: A-60916/RAO  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: (415) 494-8700  
TELEFAX: (415) 494-8771  
TELEX: 910 277299  
INFORMATION FOR SEQ ID NO: 1:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 1441 base pairs  
TYPE: nucleic acid  
STRANDEDNESS: double  
TOPOLOGY: linear  
MOLECULE TYPE: cDNA  
FEATURE:  
NAME/KEY: CDS  
LOCATION: 49..984  
US-08-414-625-1  
Query Match 100.0%; Score 1435; DB 1; Length 1441;  
Best Local Similarity 100.0%; Pred. No. 5e-300;  
Matches 1435; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
QY 1 CTGGGCGGCGTGGGAAACCCAGGCCCCCGCGAGCGGCGAGAGGTGAGATGGAGCTGGG 60  
Db 1 CTGGGCGGCGTGGGAAACCCAGGCCCCCGCGAGCGGCGAGAGGTGAGATGGAGCTGGG 60  
QY 61 CAAATGAGGACGAAGAGTGGGTGGGCGAGCGCATACCTGTTGTGAGATCCTGCTGGAC 120  
Db 61 CAAATGAGGACGAAGAGTGGGTGGGCGAGCGCATACCTGTTGTGAGATCCTGCTGGAC 120  
QY 121 AAGTGTCTCTGTGGATGCTTACGCGCAACCCCAAGAGAGTGGTCAATGAGAGGCT 180  
Db 121 AAGTGTCTCTGTGGATGCTTACGCGCAACCCCAAGAGAGTGGTCAATGAGAGGCT 180  
QY 181 CTGCAAGCTGCTTGGGAGAGAGCGGCGGAGCCCGAGCGTGTGAGATGCTGAAGATC 240  
Db 181 CTGCAAGCTGCTTGGGAGAGAGCGGCGGAGCCCGAGCGTGTGAGATGCTGAAGATC 240  
QY 241 CACGCGAGCGACCCGAGCTGATGTCAGAGCTGCGATTTCTGCGGCGGAGCCCTGTGGC 300  
Db 241 CACGCGAGCGACCCGAGCTGATGTCAGAGCTGCGATTTCTGCGGCGGAGCCCTGTGGC 300  
QY 301 CGCTTCTCTCGCGCTTACCGGAGGCGGCGCTGCGCGCGGCTGCTGCAAGAGAGCTGGCG 360  
Db 301 CGCTTCTCTCGCGCTTACCGGAGGCGGCGCTGCGCGCGGCTGCTGCAAGAGAGCTGGCG 360  
QY 361 GCCGCGCTCGCCAGCACTCGGTGCGCGCTGCACTGAGAGCTGCGCGCGCGCGAGCGG 420

```

Db 361 GCGCGCTCCGACGACTGCTGTCGCTGCACTGAGCTGCGCGCGCGGCGGAGCGG 420
Qy 421 CTGACGCTTTCGCGGCGGCGGAGGAGCGCTGTTGAGTTGCAATCCGACGAGCGG 480
Db 421 CTGACGCTTTCGCGGCGGCGGAGGAGCGCTGTTGAGTTGCAATCCGACGAGCGG 480
Qy 481 GACCGGCTCCGCGGAGTGAAGAACTGCTGAGCTGAGAGATGCGTGCAGAAATCTGAATGC 540
Db 481 GACCGGCTCCGCGGAGTGAAGAACTGCTGAGCTGAGAGATGCGTGCAGAAATCTGAATGC 540
Qy 541 GCGTCCGCGGCGGCGGAGTGAAGAACTGCTGAGCTGAGAGATGCGTGCAGAAATCTGAATGC 600
Db 541 GCGTCCGCGGCGGCGGAGTGAAGAACTGCTGAGCTGAGAGATGCGTGCAGAAATCTGAATGC 600
Qy 601 CCTCTCTGTCGAGAGTGAAGAACTGCTGAGCTGAGAGATGCGTGCAGAAATCTGAATGC 660
Db 601 CCTCTCTGTCGAGAGTGAAGAACTGCTGAGCTGAGAGATGCGTGCAGAAATCTGAATGC 660
Qy 661 CAGGCTCAGCTGAGTGAAGAACTGCTGAGCTGAGAGATGCGTGCAGAAATCTGAATGC 720
Db 661 CAGGCTCAGCTGAGTGAAGAACTGCTGAGCTGAGAGATGCGTGCAGAAATCTGAATGC 720
Qy 721 TCTGTGAGTCTCAAAATGCGCGAGGCTGAGCTGAGAGATGCGTGCAGAAATCTGAATGC 780
Db 721 TCTGTGAGTCTCAAAATGCGCGAGGCTGAGCTGAGAGATGCGTGCAGAAATCTGAATGC 780
Qy 781 CCGGACCCGCGGCTGAGTGAAGAACTGCTGAGCTGAGAGATGCGTGCAGAAATCTGAATGC 840
Db 781 CCGGACCCGCGGCTGAGTGAAGAACTGCTGAGCTGAGAGATGCGTGCAGAAATCTGAATGC 840
Qy 841 GCGCTTCCAGCTGAGTGAAGAACTGCTGAGCTGAGAGATGCGTGCAGAAATCTGAATGC 900
Db 841 GCGCTTCCAGCTGAGTGAAGAACTGCTGAGCTGAGAGATGCGTGCAGAAATCTGAATGC 900
Qy 901 CTGCTGAGGCTGAGTGAAGAACTGCTGAGCTGAGAGATGCGTGCAGAAATCTGAATGC 960
Db 901 CTGCTGAGGCTGAGTGAAGAACTGCTGAGCTGAGAGATGCGTGCAGAAATCTGAATGC 960
Qy 961 ACCGATCCCAATGCGGCTGAGTGAAGAACTGCTGAGCTGAGAGATGCGTGCAGAAATCTGAATGC 1020
Db 961 ACCGATCCCAATGCGGCTGAGTGAAGAACTGCTGAGCTGAGAGATGCGTGCAGAAATCTGAATGC 1020
Qy 1021 ATGCGCTTAAAGGCTTCTTTCGCGGCTGAGTGAAGAACTGCTGAGCTGAGAGATGCGTGCAGAAATCTGAATGC 1080
Db 1021 ATGCGCTTAAAGGCTTCTTTCGCGGCTGAGTGAAGAACTGCTGAGCTGAGAGATGCGTGCAGAAATCTGAATGC 1080
Qy 1081 AACTCCACTTGGCTTATCTGCTGAGCTGAGTGAAGAACTGCTGAGCTGAGAGATGCGTGCAGAAATCTGAATGC 1140
Db 1081 AACTCCACTTGGCTTATCTGCTGAGCTGAGTGAAGAACTGCTGAGCTGAGAGATGCGTGCAGAAATCTGAATGC 1140
Qy 1141 CCAGACCACTGCGGCTGAGTGAAGAACTGCTGAGCTGAGAGATGCGTGCAGAAATCTGAATGC 1200
Db 1141 CCAGACCACTGCGGCTGAGTGAAGAACTGCTGAGCTGAGAGATGCGTGCAGAAATCTGAATGC 1200
Qy 1201 GTGCGGCGGCTGAGTGAAGAACTGCTGAGCTGAGAGATGCGTGCAGAAATCTGAATGC 1260
Db 1201 GTGCGGCGGCTGAGTGAAGAACTGCTGAGCTGAGAGATGCGTGCAGAAATCTGAATGC 1260
Qy 1261 CAGCAGCGCTTAACTGCTGAGTGAAGAACTGCTGAGCTGAGAGATGCGTGCAGAAATCTGAATGC 1320
Db 1261 CAGCAGCGCTTAACTGCTGAGTGAAGAACTGCTGAGCTGAGAGATGCGTGCAGAAATCTGAATGC 1320
Qy 1321 CTTGCGGCTGAGTGAAGAACTGCTGAGCTGAGAGATGCGTGCAGAAATCTGAATGC 1380
Db 1321 CTTGCGGCTGAGTGAAGAACTGCTGAGCTGAGAGATGCGTGCAGAAATCTGAATGC 1380
Qy 1381 TCACACTCAGGCTGAGTGAAGAACTGCTGAGCTGAGAGATGCGTGCAGAAATCTGAATGC 1435
Db 1381 TCACACTCAGGCTGAGTGAAGAACTGCTGAGCTGAGAGATGCGTGCAGAAATCTGAATGC 1435

```

```

US-08-414-625-3
; Sequence 3, Application US/08414625
; Patent No. 5563039
; GENERAL INFORMATION:
; APPLICANT: Goeddel, David V.
; APPLICANT: Hsu, Hailing
; TITLE OF INVENTION: INTRACELLULAR SIGNALING PROTEINS AND
; TITLE OF INVENTION: METHODS OF USE
; NUMBER OF SEQUENCES: 5
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: FLHR, HOEBCHE, TEST, ALBRITTON & HERBERT
; STREET: 4 Embarcadero Center, Suite 3400
; CITY: San Francisco
; STATE: California
; COUNTRY: USA
; ZIP: 94111-4187
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/414,625
; FILING DATE:
; CLASSIFICATION: 435
; ATTORNEY/AGENT INFORMATION:
; NAME: Osman, Richard A
; REGISTRATION NUMBER: 36,627
; REFERENCE/DOCKET NUMBER: A-60916/RAO
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (415) 494-8700
; TELEFAX: (415) 494-8771
; TELEX: 910 277299
; INFORMATION FOR SEQ ID NO: 3:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 1384 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: double
; TOPOLOGY: linear
; MOLECULE TYPE: cDNA
; FEATURE:
; NAME/KEY: CDS
; LOCATION: 1..681
; FEATURE:
; NAME/KEY: mat_peptide
; LOCATION: 1..681
; US-08-414-625-3
Query Match 29.8%; Score 427.6; DB 1; Length 1384;
Best Local Similarity 78.5%; Pred. No. 1.9e-83;
Matches 540; Conservative 0; Mismatches 139; Indels 9; Gaps 2;
Qy 302 GCTTCTCCGCGCTTACCGGAGGCGGCTGCGCGCGCTGACAGAGAGCTGCGG 361
Db 2 GCTTCTCCAAAGCTTACCGGAGGCGGCTGCGCACCGCGTGCAGAGGATGCGCC 61
Qy 362 CCGGCTCGCCCAAGTACGCTGCTGCTGCACTGAGAGTGCAGGCGCGCGGCGGAGGCG 421
Db 62 CCGGCTTGGCCAGAGACCGCTGCTGCTGCACTGAGAGTGCAGGCGCGGCGGAGGAGC 121
Qy 422 TGACGCTTGTGCGGAGAGAGAGGCTGTTGAGTTGATCTTACCTTACGAGAGCGG 481
Db 122 TGACAGTTGGCTGACGAGAGAGGCTGTTGAGTTGATCTTACCTTACGAGAGCGG 181
Qy 482 ACCGCTCCGAGTGAAGAACTGCTGAGCTGAGAGATGCGTGCAGAAATCTGAATGC 541
Db 182 ACCGCTCAGGAGAGAGAACTGCGGAGCTGAGAGATGAGCTGCAAACTGACGTGTG 241
Qy 542 GCTCGGCGGCGCGGAGTGAAGAGGAGGCTGCTGCGCGCGCTTTCAGAGCGCGGCTGC 601
Db 242 ACTGCACTGCGCAGGAGTGAAGCAATACAGGATGCTTTCAGAGGTTCCAGGTTTCC 301
Qy 602 CTTCTGTGCGAGTGAAGCGCGCGCGCGCGCACCTGCGCAGACTTTCGTTCC 661

```

Db 302 CCTCTCCGACCGAGGAGAA-----ACCACTGCGCGCGCTGCGCAAGACTTTCTGTTCC 3555

Qy 662 AGGGTCAGCCTGTAGTGAATGGCGCGCTGAGCTGAAGAACAAAGACGTTCCGCGCCT 7212

Db 356 ACGGGCAAGCTGTAGTGAACCGGCGACACTCTTTCAAGACACAGACGTTGCGCGCT 4115

Qy 722 CTGTGGGCTCTCAAAATGGCGCAAGGTGGGCGCTCACTGCAAGAGGCTGCCGGGCGCTGC 7812

Db 416 CGGTGGGCTCAAGTGGCGCAGGGTGGGGCGCTCGCTGCAAGCTAATGTCGGGCACTGA 4759

Qy 782 GGGACCCGGCGCTGGACTTCGCTGGCTTACGAGTACGAGCGCGAGGAACTGTACGAGCAGG 8412

Db 476 GAGATCTCGCCCTCGACTCGCTGGGCTTACGAGTATGAGCTGAATGGGCTATACAGCAGG 5359

Qy 842 CCTTCCACACTCTGTGGGGCTTCTGTCAGAGCGCGAGAGGCGCGCGCCACAGCTGACAGCGC 9012

Db 536 CTTTCCACACTCTGTGGCGCGTTTACGCAAGCGAGAGGCGCGCGTGGCCACTGCAAGCGC 5959

Qy 902 TGGTGGAGAGCACTGAGGAGAACGAGCTCACACAGCTGGC---AGAGGACTTGTGTTGGGC 958

Db 596 TGGTGGAGGCGCTGAGGAGAACGAACTCACTAATGTTCTAGCAGGAGGAACTTGTGTTGGGC 6555

Qy 959 TGACCGATCCCAATGGCGGCGCTGAGCCTA 986

Db 656 AGGCGAGCCGGAATGGCGGCTTGGGCTTA 683

RESULT 4  
US-08-983-502-3  
; Sequence 3, Application US/08983502

APPLICANT: David WALLACH  
 APPLICANT: Mark P. BOLDIN  
 APPLICANT: Tanya M. GONCHAROV  
 APPLICANT: Yuri V. GOLITSV  
 TITLE OF INVENTION: MODULATORS OF THE FUNCTION OF FAS RECEPTORS  
 TITLE OF INVENTION: AND OTHER PROTEINS  
 NUMBER OF SEQUENCES: 34  
 CORRESPONDENCE ADDRESS:  
 ADDRESSEE: Brody and Neimark  
 STREET: 419 Seventh Street N.W., Ste. 300  
 CITY: Washington  
 STATE: D.C  
 COUNTRY: USA  
 ZIP: 20004  
 COMPUTER READABLE FORM:  
 MEDIUM TYPE: Floppy disk  
 COMPUTER: IBM PC compatible  
 OPERATING SYSTEM: PC-DOS/MS-DOS  
 SOFTWARE: Patentin Release  
 CURRENT APPLICATION DATA:  
 APPLICATION NUMBER: US/08/963,502  
 FILING DATE: 16-JAN-1998  
 PRIOR APPLICATION DATA:  
 APPLICATION NUMBER: PCT/US96/10521  
 FILING DATE: 14-JUN-1996  
 PRIOR APPLICATION DATA:  
 APPLICATION NUMBER: IL 114,615  
 FILING DATE: 16-JUL-1995  
 PRIOR APPLICATION DATA:  
 APPLICATION NUMBER: IL 114,966  
 FILING DATE: 17-AUG-1995  
 PRIOR APPLICATION DATA:  
 APPLICATION NUMBER: IL 115,319  
 FILING DATE: 14-SEP-1995  
 PRIOR APPLICATION DATA:  
 APPLICATION NUMBER: IL 116,588  
 FILING DATE: 27-DEC-1995  
 PRIOR APPLICATION DATA:  
 APPLICATION NUMBER: IL 117,932  
 FILING DATE: 16-APR-1996

```

1  ALCORNEY/AGENT INFORMATION:
2
3  NAME: Brown, Roger L.
4
5  REGISTRATION NUMBER: 25,618
6
7  REFERENCE/DOCKET NUMBER: WALLACH=19
8
9  TELECOMMUNICATION INFORMATION:
10
11  TELEPHONE: (202) 628-5197
12
13  TELEFAX: (202) 737-3528
14
15  INFORMATION FOR SEQ. ID NO. 3:
16
17  SEQUENCE CHARACTERISTICS:
18
19  LENGTH: 200 base pairs
20
21  TYPE: nucleic acid
22
23  STRANDEDNESS: single
24
25  TOPOLOGY: linear
26
27  MOLECULE TYPE: cDNA
28
29  US-08-963-502-3

```

Query Match	13.9%	Score 200;	DB 4;	Length 200;
Best Local Similarity	100.0%	Pred. No. 1.1e-34;		
Matches 200; Conservative	0;	Mismatches	0;	Indels 0; Gaps 0;

QY	622	CGCGCGCGCGCGCGCCACCTGCGCCCAACCTTTTCTGTTCACAGGGTCAAGCTGTGTGAAT	681
Db	1	CGCGCGCGCGCGCGCCACCTGCGCCCAACCTTTTCTGTTCACAGGGTCAAGCTGTGTGAAT	60
QY	682	CGCGCGCTGAAGCTGAAGGACCAACAGACGTTTCGCGCGCTCTGTGGTCTTCAATATGCGCG	741
Db	61	CGCGCGCTGAAGCTGAAGGACCAACAGACGTTTCGCGCGCTCTGTGGTCTTCAATATGCGCG	120
QY	742	AAGGTGGGGCGCTCACTGCAAGCGAGGCTGCCGGCGCTGCGGGACCCGGCGCTGGACTCG	801
Db	121	AAGGTGGGGCGCTCACTGCAAGCGAGGCTGCCGGCGCTGCGGGACCCGGCGCTGGACTCG	180
QY	802	CTGGGCTTACAGATGACGAGCG	821
Db	181	CTGGGCTTACAGATGACGAGCG	200

RESULT 5  
PCT-US96-10521-3  
; Sequence 3, Application PC/TUS9610521

TITLE OF INVENTION: MODULATORS OF THE FUNCTION OF FAS RECEPTORS  
 NUMBER OF INVENTION: AND OTHER PROTEINS  
 NUMBER OF SEQUENCES: 34  
 COMPUTER READABLE FORM:  
 MEDIUM TYPE: Floppy disk  
 COMPUTER: IBM PC compatible  
 OPERATING SYSTEM: PC-DOS/MS-DOS  
 SOFTWARE: PatentIn Release #1.0, Version #1.30 (EPO)  
 CURRENT APPLICATION DATA:  
 APPLICATION NUMBER: PCT/US96/10521  
 FILING DATE:  
 CLASSIFICATION:  
 PRIOR APPLICATION DATA:  
 APPLICATION NUMBER: IL 114,615  
 FILING DATE: 16-JUL-1995  
 PRIOR APPLICATION DATA:  
 APPLICATION NUMBER: IL 114,986  
 FILING DATE: 17-AUG-1995  
 PRIOR APPLICATION DATA:  
 APPLICATION NUMBER: IL 115,319  
 FILING DATE: 14-SEP-1995  
 PRIOR APPLICATION DATA:  
 APPLICATION NUMBER: IL 116,588  
 FILING DATE: 27-DEC-1995  
 PRIOR APPLICATION DATA:  
 APPLICATION NUMBER: IL 117,932  
 FILING DATE: 16-APR-1996  
 INFORMATION FOR SEQ ID NO: 3:  
 SEQUENCE CHARACTERISTICS:  
 LENGTH: 200 base pairs  
 type: nucleic acid





```

GENERAL INFORMATION:
APPLICANT: Lewis, Randolph V.
APPLICANT: Colgin, Mark
TITLE OF INVENTION: CDNs Encoding Minor Ampullate Spider
TITLE OF INVENTION: Silk Proteins
NUMBER OF SEQUENCES: 56
CORRESPONDENCE ADDRESS:
ADDRESS: Birch, Stewart, Kolasch & Birch
STREET: P.O. Box 747
CITY: Falls Church
STATE: Virginia
COUNTRY: USA
ZIP: 22040-3487
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentin Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/458,298
FILING DATE: 02-JUN-1995
CLASSIFICATION: 530
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/209,747
FILING DATE: 14-MAR-1994
ATTORNEY/AGENT INFORMATION:
NAME: Murphy Jr., Gerald M.
REGISTRATION NUMBER: 28,977
REFERENCE/DOCKET NUMBER: 1447-104P
TELECOMMUNICATION INFORMATION:
TELEPHONE: 703-205-8000
TELEFAX: 703-205-8050
INFORMATION FOR SEQ ID NO: 1:
SEQUENCE CHARACTERISTICS:
LENGTH: 2793 base pairs
TYPE: nucleic acid
STRANDEDNESS: double
MOLECULE TYPE: linear
TOPOLOGY: linear
MOLECULE TYPE: cdna
HYPOTHETICAL: NO
ORIGINAL SOURCE:
ORGANISM: Nephila clavipes
TISSUE TYPE: minor ampullate gland
FEATURE:
NAME/KEY: CDS
LOCATION: 183..2675
OTHER INFORMATION: /product= "N. clavipes minor
OTHER INFORMATION: ampullate silk protein"
US-08-458-298-1

Query Match 4.8%; Score 69.2; DB 1; Length 2793;
Best Local Similarity 44.1%; Pred. No. 2.4e-06;
Matches 383; Conservative 0; Mismatches 478; Indels 7; Gaps 2;
QY 104 TGGAGTCTCGCTGACAGAGTGTCTCTGCGATGCTTACGCGCAACCCGACAGAAAG 163
DB 1127 TGGAGTGTGTGTTAGTGAAGAGTGTGTGAGCTGGAAGCGGCTGAGAGTGC 1186
QY 164 TGGCAGTGTACAGGGCTTCTGCAAGCTCTTGACAGAGCGGCGGAGACCCGAGAGTGC 223
DB 1187 AGGACAGAGAGCTGAGGAGCTTCAAGTGTATCGGTGCGGAGAGAGAACTGC 1246
QY 224 TGCAGATGTCTGAAGATCCACCGAGAGACCCGAGCTGATCGAGCTGCGATTCTGG 283
DB 1247 TGCAGTGTGTGAGAGAGAGCTGAGAGCGCTGTGTGTTACGTTAGTGTCTGTGA 1306
QY 284 GCGCGAGCCCTGTGCGCCCTTCTCTCGCGCTTACCCGAGAGGGGCGCTGCGCGCGCGC 343
DB 1307 GAGAGAGCCGCTGCGGTCTGTGAGAGAGAGAGCGCGTGC---GGGTGAGAGAGCTGAG 1362
QY 344 TGCAGAGAGAGCTGCGCGCGCGCTGCGCCAGACTGCGTCCGCTGCAACTGAGAGTGC 403
DB 1363 CTACGATGTCTCAAGTGTGAGTGTGCTGCGATGAGAGAGCTGGTGCGCTGTGCTGG 1422

```

```

QY 404 GCGCGCGCCGAGCGCTGACCGCTTGTGCGGACGAGAGCGCTGTTTGAATTGA 463
DB 1423 AGCAGAGAGCTGAGAGCGCTGTGTTACGGTGAAGGTGTGTCTGAGAGCTGAC 1482
QY 464 TCTTAGCCCAAGACCCGAGCCGCTCCGGGATGAAAGACTGTGCTGAGTGAAGATCCG 523
DB 1483 TGCAGGCGCAGAGACTGAGAGCTTACGGTGTCAAGGTGATACGGTCCGAGAGAG 1542
QY 524 TGCAGAAATCTGAATGCGGCTCGGGAGCCCGGGGTGCGACGCGGAGATCGCTTCG 583
DB 1543 TGTGCTGTCTCAGCTGTCTGTAACAGAGCCGAGCGCTGTGTGTTACGTAGAGTGC 1602
QY 584 CTTGCAAGCCCGGCTGCTCTGTGCGAGTGAAGCCGCGCGCGCGCCGCTG 643
DB 1603 TGTGCTGTGAGCTGTGCTGCTGCGGGGAGGTGAGGACCCGCTGTGTGATATGG 1662
QY 644 CCCAGACTTTTCTGTTCAGAGGTACGCTGTATGAAATGGCCCTGAGCTGAAGAGAC 703
DB 1663 TGCACAGCGGTTATGTGTGCGGAGAGAGCTGTGCGGCTGTCTGTGTGACAG 1722
QY 704 AACAGAGTTGCGCGCTCTGTGAGTCTCAATAGCGCAAGTGGGCGCTCACTGCAGC 763
DB 1723 AGCAGAGAGTCTGTGTACGTTAGAGTGTGTGTGAGCTGTGAGCTGTGAGAGTGC 1782
QY 764 GAGGCTGCGGCGCTGTGCGGAGCCCGGCGCTGAGCTGCTGAGCTTACGAGAGCGCG 823
DB 1783 TGAAGCTGAGAGCCGCTGTGAGGTGAGAGAGAGAGCTGAGAGCTTACGAGTGTG 1842
QY 824 AGGAGCTGTACAGAGAGCTTCCAGCTGTGTGCGCGCTTGTGTACAGGC---GAGGCC 880
DB 1843 ATACGCTGCCGAGCAAGAGCTGTGTGCTGCGGAGCTGTGAGCAGAGAGCTGAGAG 1902
QY 881 GCGCGGCGCAGCTGAGAGCGCTGTGTGAGGCACTGAGAGAGAGAGCTCACAGCTG 940
DB 1903 TCGGATTACAGTGAAGTGTGTGTGAGAGCGGCTGTGTGTGAGTGTGAGCGCGCTGC 1962
QY 941 CAGAGAGCTTGTGCGCTGACCGATCC 968
DB 1963 AGGTGACAGAGAGAGCTGAGAGCTAC 1990

RESULT 8
US-09-249-585A-2
; Sequence 2, Application US/09249585A
; Patent No. 6417002
; GENERAL INFORMATION:
; APPLICANT: Horlick, Robert
; TITLE OF INVENTION: METHOD FOR MAINTENANCE AND SELECTION OF EPISOMES
; FILE REFERENCE: 0867/0D905
; CURRENT APPLICATION NUMBER: US/09/249,585A
; CURRENT FILING DATE: 1999-02-11
; NUMBER OF SEQ ID NOS: 18
; SOFTWARE: Patentin version 3.0
; SEQ ID NO 2
; LENGTH: 1926
; TYPE: DNA
; ORGANISM: Epstein Barr Virus
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (1)..(1926)
; OTHER INFORMATION: coding strand of EBNA-1 DNA
US-09-249-585A-2

Query Match 4.4%; Score 63.8; DB 4; Length 1926;
Best Local Similarity 42.8%; Pred. No. 3.3e-05;
Matches 320; Conservative 0; Mismatches 427; Indels 0; Gaps 0;
QY 103 GTGAGTCTCTGCTGACAGAGTGTCTGTGATGCTTACGCGCAACCCGACAGAGAG 162
DB 260 GTGAAACAGAGACAGAGACAGAGAGCGGAGGCGCAGAGAGAGGAGGAGAGAGAG 319
QY 163 GTGCAAGTGTACAGGCTTCTGAGAGCTGCTTGGCAGAGAGCGGCGGAGAGCCGAG 222

```

D	b		320	GAGGGGCAGGACGAGAGGAAGAGGGCCAGAGAGGGGCGAGAGAGGGGCGAGAGCGC	379
Q	y	223	CTGCATGATCTGAAAGATCAACCGACGCAACCGCAGTGAATGTGCAGTTCTGC	2822	
D	b	380	GAGGAGGGGCGAGAGCAGAGAGAGAGGGGCGCAGAGGGGCGAGAGCGGCGAGAGAGAG	4339	
Q	y	283	GGGGCGGCAGCCCTGTGGCCGCTTCTCTCCGCCCTAACCCGAGGGGGGCGTGCGCGCGC	3424	
D	b	440	GGGCGAGGACGAGAGGAAGAGGGGCGAGAGGGGCGAGAGCGAGAGAGGGGCGAGAGGGGCGAG	4999	
Q	y	343	CTGCAGAGGAACTTGCGCGGCGCGCTCGCCAGCACTCGGTGCGCGCTGCAACTTGAAGCTG	4020	
D	b	500	GAGGGGCGAGAGCAGAGAGAGAGGGGCGAGAGCAGAGAGAGGGGCGAGAGGGGCGAGAGCAG	5599	
Q	y	403	CGCGCCGCGCCGAGACCGGCTGGAACGCTTTGCTGCGGAGCAGAGAGCGCTGTTGAGTTGC	4624	
D	b	560	GAGGAGGGGCGAGAGGGGCGAGAGGGGCGCAGAGCAGAGAGAGGGGCGAGAGCAGAGAGGGG	6199	
Q	y	463	ATTCCTAAGCCGACAGACCCGCGCTCCGGGATTAAAGACTGACTGAGCTGAGAGATGCG	5222	
D	b	620	CAGAGAGGGGCGAGAGGGGCGAGAGCAGAGAGGGGCGCAGAGCGAGAGAGGGGCGAGAGGGG	6799	
Q	y	523	CTGCGAATACTGAAAGTGCGGCTCGGGGGGCCGCGGGGTGCGACGCGGGAGAGGTGCTTGGGCC	5822	
D	b	680	CAGGAGGGGCGAGAGCAGAGAGGGGCGAGAGCAGAGAGGGGCGAGAGAGGGGCGAGAG	7399	
Q	y	583	CCCTTGCAGCCCCCGGTGCGCTCTCTGTGAGGTGAAGCCGCGCGCGCGCCACCT	6424	
D	b	740	CAGGAGGGGCGAGAGGGGCGAGAGCAGAGAGGGGCGCAGAGAGGGGCGAGAGAGGGGCGAG	7999	
Q	y	643	GCCCAAGCTTTTCTGTTCAGAGGTCAGCTGTAGTAATGCGCCGCTGAGCCTGAAGAAGAC	7020	
D	b	800	GAGGGGCGAGGACGAGAGGAAGAGGGGCGAGAGCGGCGCAGAGCGAGAGAGGGGCGAGAGGGG	8599	
Q	y	703	CAACAGACGTTGCGGCGGCTGTGAGGTCTCAAATGGGCGAAGTGGGGCCCTCATCTGACAG	7622	
D	b	860	CAGGACGAGAGGGGCGAGAGAGGGGCGAGAGCAGAGAGGGGCGAGAGGGGCGAGAGCAGAG	9199	
Q	y	763	CGAGGCTGCCCCGGGCGTGGCGGGAACCCCGCGCTGGACTTCGCTGCGCTTAGTACGAGCGC	8222	
D	b	920	GGGCGAGAGGGGCGAGAGCAGAGAGAGGGGCGCAGAGCAGAGAGGGGCGAGAGAGGTG	9799	
Q	y	823	GAGGAGCTGTACGAGCAGAGCCTTCCAG	849	
D	b	980	GAGGCCGGGCTCGAGAGGCGAGTGAAG	1006	

RESULT 9  
US-09-050-863-2  
Sequence 2, Application US/09050863  
Patent No. 611411  
GENERAL INFORMATION:  
APPLICANT: Lao, Ying  
APPLICANT: Hiang, Betty  
TITLE OF INVENTION: Mammalian Protein Interaction Cloning  
NUMBER OF SEQUENCES: 5  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: Flehm, Honbach, Test, Albritton & Herbert  
STREET: 4 Embarcadero Center, Suite 3400  
CITY: San Francisco  
STATE: CA  
COUNTRY: USA  
ZIP: 94111-4187  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: PatentIn Release #1.0, Version #1.30  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/09/050\_863

	FILING DATE:	30-MAR-1998	
	CLASSIFICATION:		
	ATTORNEY/AGENT INFORMATION:		
	NAME:	Silva, Robin M.	
	REGISTRATION NUMBER:	38,304	
	REFERENCE/DOCKET NUMBER:	A-65638/DJB/RMS	
	TELECOMMUNICATION INFORMATION:		
	TELEPHONE:	(415) 781-1989	
	TELEFAX:	(415) 949-8711	
	INFORMATION FOR SEQ ID NO:	2:	
	SEQUENCE CHARACTERISTICS:		
	LENGTH:	2580 base pairs	
	TYPE:	nucleic acid	
	STRANDEDNESS:	unknown	
	TOPOLOGY:	unknown	
	MOLECULE TYPE:	DNA	
	us-09-050-863-2		
	Query Match	4.4%; Score 63.8; DB 3; Length 2580;	
	Best Local Similarity	42.8%; Pred. No. 3,56-05;	
	Matches 320; Conservative	0; Mismatches 427; Indels	0; Gaps
Qy	103	GTTGGAGTCTCGTCGGAACAAGATGTCTCTGTCGGATGCTTCCGCACCCACGACAGAAG	162
Db	643	GTGAAACAGAGACCAGACAG	702
Qy	163	GTGCGCATGTACAGAGGCTCTGTCAGAGCTTGCTTTGGCAGAGACCGCGCGGACCCGACGTG	222
Db	703	GAGGGGCAAGAGACAG	762
Qy	223	CTGCAGATGCTGAAGATCACCGACGACCCGACGACTGATGTCAGCTTGCAATTCTGC	282
Db	763	GAGGAGGGGCGAGAGACAG	822
Qy	283	GGGGGCGAGCCCTGTGCGCGCTTCTCTCCGCCCTAACCGGAGGGGCGCTCGCGCCCG	342
Db	823	GGGCAAGACAG	882
Qy	343	CTGCAGAGAGACCTTGACGGCCCGCTCGCCACGACACTTGTCGCGCTTCAACTGAGCTG	402
Db	883	GAGGGGCAAGAGACAG	942
Qy	403	CGCGCCGCGCGAGAGCGGCTGGAACGCTTTGCTGCGGACGAGAGACGCTGTTGAGTTGC	462
Db	943	GAGGAGGGGCGAGAGAGGGGCAAGAGGGGCAAGAGCAGAGAGAGAGAGAGAGAGAGAG	1002
Qy	463	ATCTCAATGCCAGACAGCCCGACCGGCTCCGGGATGAAGAATCTGCTGACTGAGGATGCG	522
Db	1003	CAGAGAGGGGCGAGAGAGGGGCAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG	1062
Qy	523	CTGCGAATACTGAAGTGGGCGCTCGGGGGGCGCGGGATGGCGACGGGAGAGTGGCTTCCGCC	582
Db	1063	CAGAGAGGGGCGAGAGACAGAGAGGGGCAAGAGAGAGAGAGAGAGAGAGAGAGAGAG	1122
Qy	583	CCCTTGACAGCCCCCGGTGCTCTCTGTGCGAGGTGAACCGCGCGCGCGCCGACCACT	642
Db	1123	CAGAGAGGGGCGAGAGAGGGGCAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG	1182
Qy	643	GCCCAAGACTTTTCTGTTCTCAAGGGTCAAGCTTGAATGGAATCGGCGGCTGAGCTTAAGAAC	702
Db	1183	GAGGGGCGAGAGACAG	1242
Qy	703	CAACAGACGTTCCGCGCGCTCTGTGGGTCTCAATATGGCCCAAGATGGGGCGGCTCACTGCAG	762
Db	1243	CAGAGACAGAGAGGGGCGAGAGAGGGGCAAGAGACAGAGAGGGGCGAGAGAGAGAGAG	1302
Qy	763	CGAGGCTGCGGGGCGCTGCGGAGACCCGCGCTGGAATCTGCTGGCTTAGAGTACGAGCGC	822
Db	1303	GGGGCAAGAGGGGCGAGAGACAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG	1362
Qy	823	GAGGACCTGTACAGAGAGAGGCTTCCAG	849
Db	1363	GAGGCGGGGCTCGAGAGGCAGTGAAG	1389

RESULT 10  
US-09-359-081-2  
Sequence 2, Application US/09359081  
Patent No. 6316223  
GENERAL INFORMATION:  
APPLICANT: Lao, Ying  
Huang, Betty  
Payan, Don  
TITLE OF INVENTION: Mammalian Protein Interaction Cloning  
System  
NUMBER OF SEQUENCES: 5  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: Flehr, Hohbach, Test, Albritton & Herbert  
STREET: 4 Embarcadero Center, Suite 3400  
CITY: San Francisco  
STATE: CA  
COUNTRY: USA  
ZIP: 94111-4187  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: Patent Release #1.0, Version #1.30  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/09/359,081  
FILING DATE: 22-Jul-1999  
CLASSIFICATION: <Unknown>  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: 09/050,863  
FILING DATE: <Unknown>  
ATTORNEY/AGENT INFORMATION:  
NAME: Silva, Robin M.  
REGISTRATION NUMBER: 38,304  
REFERENCE/DOCKET NUMBER: A-65638/DJB/RMS  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: (415) 781-1989  
TELEFAX: (415) 949-8711  
INFORMATION FOR SEQ ID NO: 2:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 2580 base pairs  
TYPE: nucleic acid  
STRANDEDNESS: unknown  
TOPOLOGY: unknown  
MOLECULE TYPE: DNA  
SEQUENCE DESCRIPTION: SEQ ID NO: 2:  
US-09-359-081-2  
Query Match 4.4%; Score 63.8; DB 4; Length 2580;  
Best Local Similarity 42.8%; Pred. No. 3.5e-05;  
Matches 320; Conservative 0; Mismatches 427; Indels 0; Gaps 0;  
QY 103 GTGAGTCTCTGCTGAGCAAGGTGCTCTGTGAGATGCTTACGCGCAACCCCGACAGAA 162  
DB 643 GTGGAACAGACAG 702  
QY 163 GTGCAAGTGTACAGAGGCTTGTCAAGGCTGCTTGTGAGAGAGAGAGAGAGAGAGAGAG 222  
DB 703 GAGGGGACAGAGACAG 762  
QY 223 CTGCAAGATGCTGAAGATTCACCGACAGACCGCAGACTATGATGATGATGATGATGAT 282  
DB 763 GAGGAGGGGACAG 822  
QY 283 GGGCGGACGCTGTGTGCGCTTCTCTCGGCTTACCGGAGAGAGAGAGAGAGAGAGAGAG 342  
DB 823 GGGCAGAGACAG 882  
QY 343 CTGCAAG 402  
DB 883 GAGGGGACAG 942

QY 403 CGCCCGGCGCGAGCGGCTGTGACGCTTTGCTGCGGACAGAGAGCCCTGTTGAGTTGC 462  
DB 943 GAGAGAGGGGACAG 1002  
QY 463 ATCTTACCCCTGACAG 522  
DB 1003 CAGAGAGGGGACAG 1062  
QY 523 CTGCGAAATCTGAAGTGTGCTGCGGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 582  
DB 1063 CAGAGAGGGGACAG 1122  
QY 583 CCTTTCAG 642  
DB 1123 CAGAGAGGGGACAG 1182  
QY 643 GCCAGATTTTCTGTTTCCAG 702  
DB 1183 GAGGGGACAG 1242  
QY 703 CAACAGAGCTTCGCGCGCTCTGTGAGGCTTCAATGCGCAAGGTGAGGAGCTTCACTGAC 762  
DB 1243 CAG 1302  
QY 763 CGAGGCTGCGGAG 822  
DB 1303 GGGCAG 1362  
QY 823 GAGGAGCTGTACAG 849  
DB 1363 GAGGCGGGGCTGAG 1389  
RESULT 11  
US-09-130-114-1/c  
Sequence 1, Application US/09130114  
Patent No. 5976807  
GENERAL INFORMATION:  
APPLICANT: Horlick, Robert A.  
APPLICANT: Damaj, Basam B.  
APPLICANT: Robbins, Alan K.  
TITLE OF INVENTION: Eukaryotic Cells Stably Expressing Genes  
TITLE REFERENCE: 0867/1D903US1  
FILE REFERENCE: 0867/1D903US1  
CURRENT APPLICATION NUMBER: US/09/130,114  
CURRENT FILING DATE: 1998-08-06  
NUMBER OF SEQ ID NOS: 36  
SOFTWARE: FastSeq for Windows Version 3.0  
SEQ ID NO: 1  
LENGTH: 5452  
TYPE: DNA  
ORGANISM: VEBNA  
US-09-130-114-1  
Query Match 4.4%; Score 63.8; DB 2; Length 5452;  
Best Local Similarity 42.8%; Pred. No. 4e-05;  
Matches 320; Conservative 0; Mismatches 427; Indels 0; Gaps 0;  
QY 103 GTGAGTCTCTGCTGAGCAAGGTGCTCTGTGAGATGCTTACGCGCAACCCCGACAGAA 162  
DB 2162 GTGGAACAGACAGAGACAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 2103  
QY 163 GTGCAAGTGTACAGAGGCTTGTCAAGGCTGCTTGTGAGAGAGAGAGAGAGAGAGAGAG 222  
DB 2102 GAGGGGACAG 2043  
QY 223 CTGCAAGATGCTGAAGATTCACCGACAGACCGCAGACTGATGTCAGCTGTCGATTCTGC 282  
DB 2042 GAGGAGGGGACAGAGACAG 1983  
QY 283 GGGCGGACGCTGTGTGCGCTTCTCTCGGCTTACCGGAGAGAGAGAGAGAGAGAGAGAG 342  
DB 1982 GGGCAGAGACAG 1923

QY 343 CTGCAGAGAGCCTGCGCGCCGCTCGCCAGCACTCGGTGCGCTGCACTGAGACTG 402  
Db 1922 GAGGGGAG 1663  
QY 403 CGCGCCGCGCCGAG 462  
Db 1862 GAG 1803  
QY 463 ATCTAGCCAG 522  
Db 1802 CAG 1743  
QY 523 CTGCGAAATCTGAAGTGGCTGCGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 582  
Db 1742 CAG 1683  
QY 583 CCTTTCAG 642  
Db 1682 CAG 1623  
QY 643 GCCAGACTTTCTGTTCCAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 702  
Db 1622 GAGGGGAG 1563  
QY 703 CAAAG 762  
Db 1562 CAG 1503  
QY 763 CGAGGCTGCGGCGCTGCGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 822  
Db 1502 GAGGAG 1443  
QY 823 GAGGAG 849  
Db 1442 GAGGAG 1416

RESULT 12  
US-08-910-647-1  
Sequence 1, Application US/08910647  
Patent No. 6251433  
GENERAL INFORMATION:  
APPLICANT: Zuckermann et al.  
TITLE OF INVENTION: Compositions and Methods for  
TITLE OF INVENTION: Polynucleotide Delivery  
NUMBER OF SEQUENCES: 4  
CORRESPONDENCE ADDRESS:  
ADDRESSSEE: Chiron Corporation  
STREET: 4560 Horton Street  
CITY: Emeryville  
STATE: California  
COUNTRY: U.S.A.  
ZIP: 94608-2916  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: Patentin Releasee #1.0, Version #1.30  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/910,647  
FILING DATE:  
CLASSIFICATION: 514  
ATTORNEY/AGENT INFORMATION:  
NAME: Fujita, Sharon M.  
REGISTRATION NUMBER: 38,459  
REFERENCE/DOCKET NUMBER: 1218.002  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: (510) 923-2706  
TELEFAX: (510) 655-3542  
INFORMATION FOR SEQ ID NO: 1:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 9600 base pairs

TYPE: nucleic acid  
STRANDEDNESS: single  
TOPOLOGY: linear  
MOLECULE TYPE: DNA (genomic)  
US-08-910-647-1

Query Match 4.4%; Score 63.8; DB 4; Length 9600;  
Best Local Similarity 42.8%; Pred. No. 4.4e-05;  
Matches 320; Conservative 0; Mismatches 427; Indels 0; Gaps 0;

QY 103 GTGAGATCTCTGCTGAGCAAGTGTCTGTGAGTCTTACCGCACCACCCAGCAGAG 162  
Db 689 GTGAAACAG 748  
QY 163 GTGACAGTGTACAGAGGCTGTGAGAGCTTGTGAGAGAGAGAGAGAGAGAGAGAGAG 222  
Db 749 GAGGGGAG 808  
QY 223 CTGCAGATGCTGAAGATCCAGCCAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 282  
Db 809 GAG 868  
QY 283 GGGCGGAGAGCTGTGTGAG 342  
Db 869 GGGCAG 928  
QY 343 CTGCAG 402  
Db 929 GAGGGGAG 988  
QY 403 CGCGCCGCGCCAG 462  
Db 989 GAGGAG 1048  
QY 463 ATCTTAGCCAG 522  
Db 1049 CAG 1108  
QY 523 CTGCGAAATCTGAAGTGGGCTTGGGGGCGCCGGGAGAGAGAGAGAGAGAGAGAGAG 582  
Db 1109 CAG 1168  
QY 583 CCTTTCAG 642  
Db 1169 CAG 1228  
QY 643 GCCAGACTTTTCTGTTCCAG 702  
Db 1229 GAGGGGAG 1288  
QY 703 CAAAG 762  
Db 1289 CAG 1348  
QY 763 CGAGGCTGCGGCGCTGCGGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 822  
Db 1349 GGGCAG 1408  
QY 823 GAGGAG 849  
Db 1409 GAGGCGGAG 1435

RESULT 13  
US-09-620-925-1  
Sequence 1, Application US/09620925  
Patent No. 6468986  
GENERAL INFORMATION:  
APPLICANT: Zuckermann et al.  
TITLE OF INVENTION: Compositions and Methods for  
TITLE OF INVENTION: Polynucleotide Delivery  
NUMBER OF SEQUENCES: 4  
CORRESPONDENCE ADDRESS:

ADDRESSEE: Chiron Corporation  
STREET: 4560 Horton Street  
CITY: Emeryville  
STATE: California  
COUNTRY: U.S.A.  
ZIP: 94608-2916  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: Patent in Release #1.0, Version #1.30  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/09/620,925  
FILING DATE: 21-Jul-2000  
CLASSIFICATION: <Unknown>  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: 08/910,647  
FILING DATE: <Unknown>  
ATTORNEY/AGENT INFORMATION:  
NAME: Fujita, Sharon M.  
REGISTRATION NUMBER: 38,459  
REFERENCE/DOCKET NUMBER: 1218.002  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: (510) 923-2706  
TELEFAX: (510) 655-3542  
INFORMATION FOR SEQ ID NO: 1:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 9600 base pairs  
TYPE: nucleic acid  
STRANDEDNESS: single  
TOPOLOGY: linear  
MOLECULE TYPE: DNA (genomic)  
SEQUENCE DESCRIPTION: SEQ ID NO: 1:  
US-09-620-925-1

Query Match 4.4%; Score 63.8; DB 4; Length 9600;  
Best Local Similarity 42.8%; Pred. No. 4.4e-05;  
Matches 320; Conservative 0; Mismatches 427; Indels 0; Gaps 0;

103 GTGAGTCTCTGCTGACAGAGTGTCCTGCTGAGATGCTTACGCGCAACCCAGAGAG 162  
Db 689 GTGGAACAGAGACAG 748  
163 GTGCAAGTGTACAGAGCTCTGCAAGCTGCTTGTGCAAGAGCGCGAGAGCCCGAGCTG 222  
Db 749 GAGGGGACAGAGACAG 808  
223 CTGCAGATGCTGAAGATCCAGCCGACGACCCGACAGCTGATGTGCACTGCAATTGCG 282  
Db 809 GAGGAGGGGACAGAGACAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 868  
283 GGGCGGCGAGCCCTGTGCGCGCTTCTCCCGCGGCTACCGAGAGGGGGCGCTGCGCGCG 342  
Db 869 GGGCAGAGACAG 928  
343 CTGCAGAGAGAGCTTGGCGCGCGCTGCTGCGCCAGCACTCGATGCGCTGCACTGAGCTG 402  
Db 929 GAGGGGACAGAGACAG 988  
403 CGCGCGCGCGAGCGCTGACAGCTTTGCTGCGAGAGAGAGAGAGAGAGAGAGAGAG 462  
Db 989 GAGGAGGGGACAG 1048  
463 ATCTTACCCACAGCCGAGCCGCTCCGAGATGAGAGACTGGCTGAGCTGAGAGTGGG 522  
Db 1049 CAGAGAGGGGACAG 1108  
523 CTGCGCAATCTGAAGTGCAGGCTCGAGGGGCCCGAGGCTGGAGACCGAGAGAGTGCCTT 582  
Db 1109 CAGGAGGGGACAG 1168  
583 CCCTTGACAGCCCCCGGTGCTCTTGTGTGAGAGTGAAGCCGCGCGCCGACCACT 642

Db 1169 CAGAGAGGGGACAG 1228  
643 GCCCAGACTTTTCTGTTCCAGAGGTACCTGTAGTAATCGCCGCTGAGCTTGAAGAC 702  
Db 1229 GAGGGGACAGAGACAG 1288  
703 CAACAGAGCTTCCGCGCTCTGTGGTCTCAATGCGCAAGGTGGGGCCCTCACCTTCAG 762  
Db 1289 CAG 1348  
763 CGAGGCTGCGGGGCGTGGCGGACCCGCGCTGACTGCTGACGATACGAGCGC 822  
Db 1349 GGGCAGAGAGGGGACAGAGACAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1408  
823 GAGGAGCTGTACAGAGAGCCCTTCAG 849  
Db 1409 GAGGCCGGGCTCAGAGAGGCACTGGAG 1435

RESULT 14  
US-07-884-811-15  
Sequence 15, Application US/07884811  
Patent No. 5316921  
GENERAL INFORMATION:  
APPLICANT: Godowski, Paul J. Lokker, Nathalie A. Mart, Melanie R.  
TITLE OF INVENTION: SINGLE-CHAIN HEPATOCYTE GROWTH FACTOR VARIANTS  
NUMBER OF SEQUENCES: 21  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: Genentech, Inc.  
STREET: 460 Point San Bruno Blvd  
CITY: South San Francisco  
STATE: California  
COUNTRY: USA  
ZIP: 94080  
COMPUTER READABLE FORM:  
MEDIUM TYPE: 5.25 inch, 360 Kb floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: patin (Genentech)  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/07/884,811  
FILING DATE: 19920518  
CLASSIFICATION: 530  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER:  
FILING DATE:  
ATTORNEY/AGENT INFORMATION:  
NAME: Dreger, Ginger R.  
REGISTRATION NUMBER: 33,055  
REFERENCE/DOCKET NUMBER: 755.1  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: 415/225-3216  
TELEFAX: 415/952-9881  
TELEX: 910/371-7168  
INFORMATION FOR SEQ ID NO: 15:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 10596 bases  
TYPE: NUCLEIC ACID  
STRANDEDNESS: single  
TOPOLOGY: linear  
US-07-884-811-15

Query Match 4.4%; Score 63.8; DB 1; Length 10596;  
Best Local Similarity 42.8%; Pred. No. 4.5e-05;  
Matches 320; Conservative 0; Mismatches 427; Indels 0; Gaps 0;

103 GTGAGTCTCTGCTGACAGAGTGTCCTGTGAGATGCTTACGCGCAACCCAGAGAG 162  
Db 2223 GTGGAACAGAGACAGAGACAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 2282  
163 GTGCAAGTGTACAGAGCTCTGAGAGCTGCTTGGCAGAGAGCGCGAGAGCCCGAGACTG 222  
Db 2283 GAGGGGACAGAGACAG 2342

QY 223 CTGCAGATGCTGAAGATCCACCGCAGCACCCTGATCTGTGACGTGCGATTCTGC 282  
Db 2343 GAGAGAGGGGCGAGGAGCAGAGAGGGGCGAGAGGGGCGAGAGGAGAG 2402  
QY 283 GGGCGGAGCCCTGTGCGCTTCTCTCCGGCTTACCGCGAGGGGGCGTCCGCCGCG 342  
Db 2403 GGGGAGAGGAGCAGAGAGGGGCGAGAGGGGCGAGAGGGGCGAGAGGGGCGAG 2462  
QY 343 CTGCAGAGGAGCCCTGAGGGGCGGCTGCGCCAGCACTCGGTCGCTGCACTGAGACTG 402  
Db 2463 GAGGAGGAGGAGCAGAGAGGGGCGAGAGGAGCAGAGAGGGGCGAGAGGGGCGAG 2522  
QY 403 CGCGCCGCGCGCAGAGCGGCTGAGAGCTTGTGCGGAGCAGAGAGCGGCTTTGATTGC 462  
Db 2523 GAGAGAGGGGCGAGAGGGGCGAGAGGGGCGAGAGGGGCGAGAGGGGCGAGAGGGG 2582  
QY 463 ATCTTAGCCCAAGCAGCCCGACCGGCTCCGGGATGAAGACTGCTGAGCTGAGATGCG 522  
Db 2583 CAGAGAGGGGCGAGAGGGGCGAGAGCAGAGGGGCGAGAGGGGCGAGAGGGGCGAG 2642  
QY 523 CTGCGAATCTGAAGTGGGCTGCGGGGCGCGGGGTCGACCGGGGAGGTCGCTTGGCC 582  
Db 2643 CAGAGAGGGGCGAGAGCAGAGAGGGGCGAGAGGAGGAGGAGGAGGAGGAGGAGGAG 2702  
QY 583 CCTTTGAGCGCCCGGCTGCTCTGTGAGAGGTGAAGCGCGCGCGCGCCCACT 642  
Db 2703 CAGAGAGGGGCGAGAGGGGCGAGAGCAGAGAGGGGCGAGAGGGGCGAGAGGGGCGAG 2762  
QY 643 GCCCAACTTTTCTGTTCCAGAGGTCAAGCTGTGATGCGCCGCTGAGCTGAGAGGAC 702  
Db 2763 GAGGAGGAGGAGCAGAGAGGGGCGAGAGGGGCGAGAGGAGGAGGAGGAGGAGGAG 2822  
QY 703 CAACAGAGCTTTCGCGGCTGTGAGGTCTCAATGCGCAAGGGGCGGCTCACTGAC 762  
Db 2823 CAGAGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 2882  
QY 763 CGAGGCTGCGGGGCGTCCGAGACCCGCGCTGAGACTGCTGCTTCACTGAGAGCGC 822  
Db 2883 GGGGAGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 2942  
QY 823 GAGGAGCTGTACGAGCAGCGCTTCCAG 849  
Db 2943 GAGGCGGGGTCGAGAGGAGGAGGAG 2969

RESULT 15  
US-07-885-971-15  
Sequence 15, Application US/07885971  
Patent No. 5328837  
GENERAL INFORMATION:  
APPLICANT: Godowet, Paul J. Lokker, Nathalie A. Mark, Melanie R.  
TITLE OF INVENTION: HEPATOCYTE GROWTH FACTOR PROTEASE DOMAIN VARIANTS  
NUMBER OF SEQUENCES: 21  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: Genentech, Inc.  
STREET: 460 Point San Bruno Blvd  
CITY: South San Francisco  
STATE: California  
COUNTRY: USA  
ZIP: 94080  
COMPUTER READABLE FORM:  
MEDIUM TYPE: 5.25 inch, 360 Kb floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: patin (Genentech)  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/07/885,971  
FILING DATE: 19920518  
CLASSIFICATION: 530  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER:  
FILING DATE:

ATTORNEY/AGENT INFORMATION:  
NAME: Dieger, Ginger R.  
REGISTRATION NUMBER: 33,055  
REFERENCE/DOCKET NUMBER: 779  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: 415/225-3216  
TELEFAX: 415/952-9881  
TELEX: 910/371-7168  
INFORMATION FOR SEQ. ID NO: 15:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 10596 bases  
TYPE: NUCLEIC ACID  
STRANDEDNESS: single  
TOPOLOGY: linear  
US-07-885-971-15

Query Match 4.4%; Score 63.8; DB 1; Length 10596;  
Best Local Similarity 42.8%; Pred. No. 4.5e-05;  
Matches 320; Conservative 0; Mismatches 427; Indels 0; Gaps 0;

QY 103 GTGGAATCTGCTGAGCAAGGTGTCCTGTGCGATGCTTACGCGCAGCCCGCAGAGAG 162  
Db 2223 GTGGAACAGAGCAGAGCAGAGCGGAGCGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 2282  
QY 163 GTGCAATGTACAGAGGCTGTGACGCTGCTTGGCAGAGAGCGGCGGAGCCCGACGTG 222  
Db 2283 GAGGGGAGAGGAGCAGAGAGGGGCGAGAGGGGCGAGAGGGGCGAGAGGGGCGAGAGGAG 2342  
QY 223 CTGCAGATGCTGAAGATCCACCGCAGGACCCGAGTGTCTGCACTGCGATTCTGC 282  
Db 2343 GAGAGAGGGGCGAGAGCAGAGAGGGGCGAGAGGGGCGAGAGGGGCGAGAGGAGGAG 2402  
QY 283 GGGCGGAGCCCTGTGAGGCGCTTCTCCGCGCTTACCGGAGGGGCGGCTGGCGCGCG 342  
Db 2403 GGGGAGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 2462  
QY 343 CTGCAGAGAGCCTTGGGCGCGCTGCGCCAGCACTCGGTGCGCTGCACTGAGAGCTG 402  
Db 2463 GAGGGGAGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 2522  
QY 403 CGGCGCGGCGCGAGCGGCTGTGACGCTTGTGCGGAGGAGGAGGAGGAGGAGGAGGAG 462  
Db 2523 GAGAGAGGGGCGAGAGGGGCGAGAGGGGCGAGAGCAGAGAGGAGGAGGAGGAGGAGGAG 2582  
QY 463 ATCTTAGCCCAAGCAGCCCGACCGGCTCCGGGATGAAGACTGCTGAGCTGAGATGCG 522  
Db 2583 CAGAGAGGGGCGAGAGGGGCGAGAGGGGCGAGAGGGGCGAGAGGGGCGAGAGGGGCGAG 2642  
QY 523 CTGCGAATCTGAAGTGGGCTGTGAGGTCTCAATGCGCAAGGGGCGGCTCACTGAC 582  
Db 2643 CAGAGAGGGGCGAGAGGGGCGAGAGGGGCGAGAGGGGCGAGAGGGGCGAGAGGGGCGAG 2702  
QY 583 CCTTTGAGCGCCCGGCTGCTCTGTGAGAGGTGAAGCGCGCGCGCGCCCACT 642  
Db 2703 CAGAGAGGGGCGAGAGGGGCGAGAGGGGCGAGAGGGGCGAGAGGGGCGAGAGGGGCGAG 2762  
QY 643 GCCCAACTTTTCTGTTCCAGAGGTCAAGCTGTGATGCGCCGCTGAGACTGAGAGGAC 702  
Db 2763 GAGGAGGAGGAGCAGAGAGGGGCGAGAGGGGCGAGAGGGGCGAGAGGGGCGAGAGGGG 2822  
QY 703 CAACAGAGCTTTCGCGGCTGTGAGGTCTCAATGCGCAAGGGGCGGCTCACTGAC 762  
Db 2823 CAGAGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 2882  
QY 763 CGAGGCTGCGGGGCGTTCGAGACCCGCGCTGAGACTGCTGCTTCACTGAGAGCGC 822  
Db 2883 GGGCAGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 2942  
QY 823 GAGGAGCTGTACGAGCAGCGCTTCCAG 849  
Db 2943 GAGGCGGGGTCGAGAGGAGGAGGAGGAG 2969

Tue Feb 4 11:19:01 2003

us-09-763-748-1.rni

Page 12

Search completed: February 3, 2003, 16:14:31  
Job time : 89 secs

...

GenCore version 5.1.3  
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: February 3, 2003, 14:25:18 ; Search time 80 Seconds  
(without alignments)  
8058.688 Million cell updates/sec

Title: US-09-763-748-1  
Perfect score: 1435  
Sequence: 1 cggcgcgggcggtggaccacca.....gataataaagtataacacg 1435

Scoring table: IDENTITY\_NIC  
Gapop 10.0 , Gapext 1.0

Searched: 396772 seqs, 224632407 residues

Total number of hits satisfying chosen parameters: 793544

Minimum DB seq length: 0

Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%

Listing first 45 summaries

Database :

Published Applications NA:\*

- 1: /cgn2\_6/prodata/2/pubpna/US07\_PUBCOMB.seq:\*
- 2: /cgn2\_6/prodata/2/pubpna/US06\_NEW\_PUB.seq:\*
- 3: /cgn2\_6/prodata/2/pubpna/US06\_NEW\_PUB.seq:\*
- 4: /cgn2\_6/prodata/2/pubpna/US06\_PUBCOMB.seq:\*
- 5: /cgn2\_6/prodata/2/pubpna/US07\_NEW\_PUB.seq:\*
- 6: /cgn2\_6/prodata/2/pubpna/US07\_PUBCOMB.seq:\*
- 7: /cgn2\_6/prodata/2/pubpna/US08\_NEW\_PUB.seq:\*
- 8: /cgn2\_6/prodata/2/pubpna/US08\_PUBCOMB.seq:\*
- 9: /cgn2\_6/prodata/2/pubpna/US09\_NEW\_PUB.seq:\*
- 10: /cgn2\_6/prodata/2/pubpna/US09\_PUBCOMB.seq:\*
- 11: /cgn2\_6/prodata/2/pubpna/US10\_NEW\_PUB.seq:\*
- 12: /cgn2\_6/prodata/2/pubpna/US10\_PUBCOMB.seq:\*
- 13: /cgn2\_6/prodata/2/pubpna/US60\_NEW\_PUB.seq:\*
- 14: /cgn2\_6/prodata/2/pubpna/US60\_PUBCOMB.seq:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

#### SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	145.4	10.1	325	10	US-09-960-352-12660
2	56.2	3.9	2108	10	US-09-862-832-225
3	55.4	3.9	14800	10	US-09-854-456-1601
4	52	3.6	1156	10	US-09-827-854-11
5	51.4	3.6	1157	10	US-09-854-456-760
6	51.4	3.6	1157	10	US-09-880-107-2244
7	51.2	3.6	41907	10	US-09-967-013-5
8	50.4	3.5	1156	9	US-09-870-759-129
9	50.4	3.5	1156	10	US-09-827-854-7
10	50.4	3.5	1156	10	US-09-827-854-8
11	50.4	3.5	1156	10	US-09-827-854-10
12	50.4	3.5	1156	10	US-09-827-854-12
13	50.4	3.5	1291	12	US-10-044-090-454
14	50.4	3.5	4826	10	US-09-772-304A-1
15	50.2	3.5	88421	9	US-09-976-059-1
16	49.6	3.5	1614	9	US-09-976-740-45
17	49.6	3.5	1614	12	US-10-023-529-45
18	49.6	3.5	1614	12	US-10-023-523-45
19	49.6	3.5	12425	9	US-09-976-740-50

C	20	49.6	3.5	12425	12	US-10-023-529-50	Sequence 50, Appl
C	21	49.6	3.5	12425	12	US-10-023-523-50	Sequence 50, Appl
C	22	49.6	3.4	2176	9	US-10-032-159A-19	Sequence 19, Appl
C	23	48.8	3.4	1156	10	US-09-827-854-9	Sequence 9, Appl
C	24	48.8	3.4	2561	9	US-09-976-740-48	Sequence 48, Appl
C	25	48.8	3.4	2561	12	US-10-023-529-48	Sequence 48, Appl
C	26	48.8	3.4	2561	12	US-10-023-523-48	Sequence 48, Appl
C	27	48.4	3.4	4041	9	US-09-860-846-36	Sequence 36, Appl
C	28	48.4	3.4	4041	9	US-09-861-289-36	Sequence 36, Appl
C	29	48.4	3.4	36778	9	US-09-860-846-5	Sequence 5, Appl
C	30	48.4	3.4	36778	10	US-09-861-289-5	Sequence 5, Appl
C	31	48	3.3	516	9	US-09-991-496-34	Sequence 34, Appl
C	32	48	3.3	516	10	US-09-874-923-34	Sequence 34, Appl
C	33	48	3.3	582	9	US-09-991-496-114	Sequence 114, Appl
C	34	48	3.3	582	10	US-09-874-923-114	Sequence 114, Appl
C	35	48	3.3	7065	9	US-09-991-496-115	Sequence 115, Appl
C	36	48	3.3	7065	10	US-09-874-923-115	Sequence 115, Appl
C	37	47.8	3.3	2561	12	US-10-023-529-48	Sequence 48, Appl
C	38	47.8	3.3	2561	12	US-10-023-523-48	Sequence 48, Appl
C	39	47.8	3.3	2561	12	US-10-023-523-48	Sequence 48, Appl
C	40	47.2	3.3	1233	9	US-09-712-363-77	Sequence 77, Appl
C	41	47	3.3	3345	10	US-09-894-998-49	Sequence 49, Appl
C	42	46.6	3.2	4366	12	US-10-044-090-668	Sequence 668, Appl
C	43	45.6	3.2	11220	9	US-09-860-846-32	Sequence 32, Appl
C	44	45.6	3.2	11220	10	US-09-861-289-32	Sequence 32, Appl
C	45	45.4	3.2	457	10	US-09-960-352-12765	Sequence 12765, A

#### ALIGNMENTS

RESULT 1  
US-09-960-352-12660  
Sequence 12660, Application US/09960352  
Patent No. US20020137139A1  
GENERAL INFORMATION:  
APPLICANT: Warren, Wesley C.  
APPLICANT: Tao, Nengbing  
APPLICANT: Byatt, John C.  
APPLICANT: Mathaiagan, Nagappan  
TITLE OF INVENTION: NUCLEIC ACID AND OTHER MOLECULES ASSOCIATED WITH LACTATION AND  
FILE REFERENCE: 16511.006/37-21(10298)C  
CURRENT APPLICATION NUMBER: US/09/960,352  
CURRENT FILING DATE: 2001-09-24  
NUMBER OF SEQ ID NOS: 15112  
SEQ ID NO 12660  
LENGTH: 325  
TYPE: DNA  
ORGANISM: Bos taurus  
OTHER INFORMATION: Clone ID: 54-LIB3058-020-Q1-K1-F6  
US-09-960-352-12660

Query Match 10.1%; Score 145.4; DB 10; Length 325;  
Best local similarity 86.4%; Pred. No. 2, 7e-24;  
Matches 172; Conservative 0; Mismatches 26; Indels 1; Gaps 1;  
QY 1 CTGCGGCGGTGGGACCCAGGCGGCGGCGGAGGTGATGCGAGCTGG 60  
DB 22 CCGCGGCGGTGGGACCCAGGCGGCGGCGGCGGAGGTGATGCGAGCTGG 80  
QY 61 CAAATGCGGACGAAAGGTGGGCGGAGGCACTGTTTGGAGCTCGCTGGAC 120  
DB 81 CCAATGCGGCTTGAAGGTGGGCGGAGGCACTGTTTGGAGCTCGCTGAC 140  
QY 121 AAGGTGCTCTGTCGATGCTAGCGACCCAGGAGGAGGTGAGTACAGGCT 180  
DB 141 AAGGTGCTCTGTCGATGCTAGCGACCCAGGAGGAGGTGAGTACAGGCT 200  
QY 181 CTGCGGCTGCTTGGGAG 199  
DB 201 CTCGAGACTGCCCTAGCAG 219



```
RESULT 2
US-09-962-832-225
; Sequence 225, Application US/09962832
; Patent No. US20020110821A1
; GENERAL INFORMATION:
; APPLICANT: Ebner, Reinhard
; TITLE OF INVENTION: Cancer Gene Determination and Therapeutic Screening Using Signatu
; FILE REFERENCE: 689290-74
; CURRENT APPLICATION NUMBER: US/09/962,832
; CURRENT FILING DATE: 2001-09-25
; PRIOR APPLICATION NUMBER: US/60/235,077
; PRIOR FILING DATE: 2000-09-25
; PRIOR APPLICATION NUMBER: US/60/235,280
; PRIOR FILING DATE: 2000-09-25
; NUMBER OF SEQ ID NOS: 259
; SOFTWARE: PatentIn version 3.0
; SEQ ID NO 225
; LENGTH: 2108
; TYPE: DNA
; ORGANISM: Homo sapiens
US-09-962-832-225

Query Match          3.9%; Score 56.2; DB 10; Length 2108;
Best Local Similarity 45.0%; Pred. No. 0.0007;
Matches 303; Conservative 0; Mismatches 358; Indels 12; Gaps 2;

QY 255 GCAGCTGATCGTCAGCTGCGATTCTGCGGGCGAGCCCTGTGGCGCTTCTCCGCGC 314
    |||||
DB 620 GCAGCAGGAGGGGCGAGCTGAGCTCCACAGCAGCAGGAGGGGCGAGCTGAGCTCCACA 679
    |||||
QY 315 CTAACCGGAGGGGGGCTGCGGGCGGCTGCAAGAGAACCTGGGCGGCGCTGCGCCCA 374
    |||||
DB 680 GCAGCAGGAGGGGCGAGCTGAGCTCCACAGCAGCAGGAGGGGCGAGCTGAGCTCTGA 739
    |||||
QY 375 GCAGCTGATGCGCTGCACTGAGAGCTGCGCGCGCGCGAGCGGCTGAGAGCTTGTCT 434
    |||||
DB 740 GCA----GCAAGAGGGGCGAGCTGAGCTCTCTGAGCAGCAGAGGGGCGAGCTGAGCTCTC 796
    |||||
QY 435 GCGCGAGCAGAGCGCTGTTTGAATTGATCTTACGCCACAGCCCGACCGGCTCCGGGA 494
    |||||
DB 797 TGAGCAGCAGGAGGAGCAGCTGAAGCAGCTGAGCAGCAGGAGGGGCGAGCTGAGAGTCCC 856
    |||||
QY 495 TGAAGACCTGCTGAGCTGAGAGATGCGCTGGAATCTGAAGTGGCGCTCGGGGCGCG 554
    |||||
DB 857 AGAGGAGCAGATGGGCGAGCTGAAGTACTTGAAACAGCAGGAGGGGCGAGCTGAAGCAGCT 916
    |||||
QY 555 GGGTGGCGAGCGGAGAGTGTCTTGGCCCTTGGCAGCCCGGTTGCCCTCTCTGTGGA 614
    |||||
DB 917 GGATCAGCAGAGGAAGCAGCCAGAGGTCTCCAGAGCAGAGATGGGGCAGCTGAAGCAGCT 976
    |||||
QY 615 GGTGAAGCCGCGCGCGCGCCACCTGCGCAACTTTTCTGTTTCAGAGTTCAGCTGT 674
    |||||
DB 977 GGAGGAGCAGGAGGGGCGAGCTTAAGCATCTGAGCAGCAGAGGGGCGCACTGAAGCAGCT 1036
    |||||
QY 675 AGTGAATGGCGCGCTGAGCTGAAGGACCAAGACGTTTGGCGCTGTGTGGTCTCA 734
    |||||
DB 1037 GAGGAGCAGGAGGGGCGAGCTGAAGCAGCTGAGCAGCAGGAGGGGCGAGCTGAGAGCTT 1096
    |||||
QY 735 ATGCGCAGAGTGGGGCGCTCACTGAGCAGAGCTGCGGGGCGCTGCGGAGCCCGGCGCT 794
    |||||
DB 1097 GGAGGAGCAGGAGG-----GAGGCTGGGGGCTCCAGAGCAGCAGAGTGTGTGAGCT 1147
    |||||
QY 795 GGAATCGCTGCGCTTACAGTACGAGCGCAGGGGAGCTGTACGAGCAGGCGCTTCAAGTGTCT 854
    |||||
DB 1148 GAAAGCAGCTAGAGAGCAGCAGGGGCGAGCAGCAAGCAGCTGAGAGAGAGAGAGGGGAGCT 1207
    |||||
QY 855 GCGGCGCTTCTGTGAGAGCGGAGAGGGCGCGCGCCAGCGCTGACGCGCTGTGTGAAGCAGCT 914
    |||||
DB 1208 GAAAGCAGCTGTGAGCAGCAGGAGGGGCGAGCTGAAGCATCTGTGTGAGCAGCAGGAGGGGAGCT 1267
    |||||
QY 915 CGAGAGAGAGAG 927
    |||||
```

```
DB 1268 GGAGCAGCAGAG 1280
    |||||
RESULT 3
US-09-954-456-1601
; Sequence 1601, Application US/09954456
; Patent No. US20020115057A1
; GENERAL INFORMATION:
; APPLICANT: Young, Paul
; TITLE OF INVENTION: Process for Identifying Anti-Cancer Therapeutic Agents Using Canc
; FILE REFERENCE: 689290-76
; CURRENT APPLICATION NUMBER: US/09/954,456
; CURRENT FILING DATE: 2001-09-18
; PRIOR APPLICATION NUMBER: US/60/233,617
; PRIOR FILING DATE: 2000-09-18
; PRIOR APPLICATION NUMBER: US/60/234,052
; PRIOR FILING DATE: 2000-09-20
; PRIOR APPLICATION NUMBER: US/60/234,923
; PRIOR FILING DATE: 2000-09-25
; PRIOR APPLICATION NUMBER: US/60/235,134
; PRIOR FILING DATE: 2000-09-25
; PRIOR APPLICATION NUMBER: US/60/235,637
; PRIOR FILING DATE: 2000-09-26
; PRIOR APPLICATION NUMBER: US/60/235,638
; PRIOR FILING DATE: 2000-09-26
; PRIOR APPLICATION NUMBER: US/60/235,711
; PRIOR FILING DATE: 2000-09-27
; PRIOR APPLICATION NUMBER: US/60/235,720
; PRIOR FILING DATE: 2000-09-27
; PRIOR APPLICATION NUMBER: US/60/235,840
; PRIOR FILING DATE: 2000-09-27
; PRIOR APPLICATION NUMBER: US/60/235,863
; PRIOR FILING DATE: 2000-09-27
; NUMBER OF SEQ ID NOS: 2276
; SOFTWARE: PatentIn version 3.0
; SEQ ID NO 1601
; LENGTH: 14800
; TYPE: DNA
; ORGANISM: Homo sapiens
US-09-954-456-1601

Query Match          3.9%; Score 55.4; DB 10; Length 14800;
Best Local Similarity 43.0%; Pred. No. 0.0017;
Matches 382; Conservative 0; Mismatches 501; Indels 6; Gaps 2;

QY 30 GAGGCGGCGCAGAGGTGAGATGCGAGCTGGGCAGAAATGGGACAGAAAGTGGTGGGAG 89
    |||||
DB 4662 GCGGCGAGGCGGAGGTGAGAGCTGGGCTCGCGGTGAAGGCCGAGACCGAGCGCGCGCA 4721
    |||||
QY 90 CGCATACTCTGTTTGTGAGATCTCTGCTGACAGAGTGTCTCTGTGAGATGCTTACGCCCA 149
    |||||
DB 4722 GAAGCAGAGGGGCGCTGCGAGGCGCTGGAAGCTGGCGCTGAGGCGGAGAGCGGACCG 4781
    |||||
QY 150 CCCCAGAGAGAGGTGAGAGTACAGGGCTCTGAGAGGCTGCGCTTGGC---AGAGAGCGG 206
    |||||
DB 4782 GCGCTTGGCGGAGGCGCGAGGTGAGACGAGCCGCGAGGTACAGGTGAGGCTCTGAGAGCGG 4841
    |||||
QY 207 CCGAGCCCGGACGTGCTGAGATGCTGAAGATCCACCGCAGCAGCAGCCGAGCTGATGCT 266
    |||||
DB 4842 GCAGGCGAGTGAAGAGGCGGAGGCTGCAAGACAAAGCGCTCTTCCGCGAGAAAGAGCGG 4901
    |||||
QY 267 GCAGCTGAGATTTCTGGGGGCGCAGCCCTTGTGGCGGCTTCTCCGCGGCTTACCGGAGGG 326
    |||||
DB 4902 ACAGCTGAGAGGCTCTCTGCGAGAGGAAACAGTGGCTGTGCAACAGCTGCGGAGAGAGGC 4961
    |||||
QY 327 GCGCGTGGCGCGCGCGCTGCGAGAGAGCCTGAGCGGCGCGCTGCGCCAGCATCGATGCC 386
    |||||
DB 4962 TGAAGCGGCGGACAGCAGCAGGCGCGAGGCGCGAGGCGGCGGCGAGAGAGAGCGGA 5021
    |||||
QY 387 GCTGCACTGAGAGCTGCGCGCGCGCGCGAGCGGCTGAGACGCTTGTGCTGCGGAGAGGA 446
    |||||
```

Db 5022 GCTGAGCGCTGAGCTCAAGGCCAACGAGGCGCTACGCTGCGCTGACGAGCGGAG 5081  
Qy 447 GCCTGTTTGAAGTTGATCTTACGCGCCAGACCGGCTCCGGGATGAAGAACTGAC 506  
Db 5082 GGTGGCCGACGAGAGAGCTTGGC---GCAAGCCGAGGCTTAAGAACGAGAGAGGCT 5138  
Qy 507 TGAAGCTGAGAGTACGCTGCGGAAATCTGAAGTGCCTGCGGAGGCGCGGAGTGCAGCG 566  
Db 5139 GGAAGCGGAGCGGCGCGCGCGGAGGCGGAGAGAGAGGCGCTGCGGAGCGGAGCT 5198  
Qy 567 GGAGTGCCTTCCGCCCCCTTTCAGACCCCGGCTCCTCTCTGTGCGAGTGAAGCCGCC 626  
Db 5199 GGCTGAACAGAGCTGAGAGAGAGGCGGACGCTGCGGAGAGCAACCGGCGAGCGCTCT 5258  
Qy 627 GCGCGCGCGCGCACCTGCGCGACCTTTCTGTCCAGGGGTGAGGCTGTAAGTAATCGGCG 686  
Db 5259 GCGCGCGGAGAGAGTGTGATTCGCGCTGCGGCGGAGAGCGAGGCGGAGAGCGG 5318  
Qy 687 GCTGAGCTTGAAGAACCAAGAGCTTTCGCGCTCTGTGAGTCTCAATGAGCGAAGGT 746  
Db 5319 GCAGCTGCTGAGAGAGAGCTGCGCGCGCTGAGGCTGAGGCGGCTGACGCGAGAA 5378  
Qy 747 GGGCGCTCATCTGACGAGCGAGCTGCGGCGCTGCGGAGACCGCGCTGAGCTGCTGCG 806  
Db 5379 ACGGACAGAGCTGAGAGCGGAGCTGCGGCGGAGAGTGAAGTGAAGTGTGCTGCGC 5438  
Qy 807 CTACGAGTACGAGCGCGGAGGAGCTGTACGACAGGCGCTTCCAGGCTGTGCGCGCTTGT 866  
Db 5439 CAGCAAGCGGAGGCTGAGAGAGTCCGCTCCACCGAGCGAATTCAGAGCAGAGGCT 5498  
Qy 867 GCAGGCGGAGGCGCGCGCGCGCGCGCTGCGCTGAGGAGGCACTC 915  
Db 5499 GGAAGCGGAGGCGCGCGCTTCCGAGAGCTGCGGAGAGGCGCGCGCGC 5547

RESULT 4  
US-09-827-854-11  
; Sequence 11, Application US/09827854  
; Patent No. US20020123093A1  
; GENERAL INFORMATION:  
; APPLICANT: Zambis, Vassilis  
; APPLICANT: Kyriacos, Kyriacos E.  
; TITLE OF INVENTION: Compounds and methods for lowering  
; FILE REFERENCE: 07180/004003  
; CURRENT APPLICATION NUMBER: US/09/827,854  
; CURRENT FILING DATE: 2001-04-05  
; PRIOR APPLICATION NUMBER: US 09/679,088  
; PRIOR FILING DATE: 2000-10-04  
; PRIOR APPLICATION NUMBER: US 09/544,386  
; PRIOR FILING DATE: 2000-04-06  
; NUMBER OF SEQ ID NOS: 19  
; SOFTWARE: FastSeq for Windows Version 4.0  
; SEQ ID NO 11  
; LENGTH: 1156  
; TYPE: DNA  
; ORGANISM: Homo sapiens  
US-09-827-854-11

Query Match 3.6%; Score 52; DB 10; Length 1156;  
Best Local Similarity 44.2%; Pred. No. 0.0053;  
Matches 214; Conservative 0; Mismatches 270; Indels 0; Gaps 0;

Qy 108 GTCTCTGCTGACAAAGGTGCTCTGCGATGCTTACGCGGCAACCCCGACGAGAGGTGCG 167  
Db 360 GACCCCGGTGGCGGAGGAGACGCGGCGACGCTGTCTCAAGAGAGCTGACGCGCGCACAGC 419  
Qy 168 AGTGTACAGGAGCTGCGAGGCTGCTTGGACAGAGAGCGCGGAGACCGGAGCTGTGCA 227  
Db 420 CCGGCTGGGCGGAGCATGAGAGAGCTGTGCGCGCTCTGTGTGAGTACCGCGGAGAGT 479  
Qy 228 GATGCTGAAGTCCACCGCAGGAGACCGCAGCTGATCTGTGACGCTGCGATTTCTGCGGCG 287

Db 480 GCAGGCCATGCTCGGCCAGAGACCGAGAGCTGCGGAGCGCTTCGCTCCACCTGCG 539  
Qy 288 GCAGCCCTGTGAGCGCTTCTCTCCGCGGCTTACCGGAGAGGAGCGCTGCGCGCGCTGCA 347  
Db 540 CAGCTGTGTAGCGGCTTCTCTCCGATGCGATGATCTGACAGAGCGCTTGGCAGTGT 599  
Qy 348 GAGAGGCTTGGCGCGCGCTCGGCCAGCACTGCTGCGCTGCACTGAGACTGCGCGC 407  
Db 600 CAGGCGCGGCGCGGAGGCGCGGAGCGCGGCTCAGGCGCATCCGCGAGGCGCTGCGG 659  
Qy 408 CCGCGCGGAGCGCTGAGCGCTTGTGCGGAGAGAGCGCTGTTGATGATCTCT 467  
Db 660 GCGCTGTGTAACAGGCGCGCGCTGCGGCGCGCATCTGTGCTCTGCGCGCGCGC 719  
Qy 468 AGCCAGACGCGGACCGGCTCCGGGATGAAGACTGCTGAGCTGAGAGATGCGCTGCG 527  
Db 720 GCTACAGAGAGGCGCGGAGCTGCGGCGGAGCGGCTGCGCGGATGAGAGAGATGG 779  
Qy 528 AAATCTGAAGTGCCTGCGGAGCGCGGAGTGCAGCGGAGGTGCTTTCGCGCGCTT 587  
Db 780 CAGCGGAGCGCGGAGCGCGCTGAGCGAGTGAAGAGCAGAGTGGCGGAGTGGCGGCA 839  
Qy 588 GCAG 591  
Db 840 GCTG 843

RESULT 5  
US-09-954-456-760  
; Sequence 760, Application US/09954456  
; Patent No. US20020115057A1  
; GENERAL INFORMATION:  
; APPLICANT: Young, Paul  
; TITLE OF INVENTION: Process for Identifying Anti-Cancer Therapeutic Agents Using Can  
; FILE REFERENCE: 689290-76  
; CURRENT APPLICATION NUMBER: US/09/954,456  
; CURRENT FILING DATE: 2001-09-18  
; PRIOR APPLICATION NUMBER: US/60/233,617  
; PRIOR FILING DATE: 2000-09-18  
; PRIOR APPLICATION NUMBER: US/60/234,052  
; PRIOR FILING DATE: 2000-09-20  
; PRIOR APPLICATION NUMBER: US/60/234,923  
; PRIOR FILING DATE: 2000-09-25  
; PRIOR APPLICATION NUMBER: US/60/235,134  
; PRIOR FILING DATE: 2000-09-25  
; PRIOR APPLICATION NUMBER: US/60/235,637  
; PRIOR FILING DATE: 2000-09-26  
; PRIOR APPLICATION NUMBER: US/60/235,638  
; PRIOR FILING DATE: 2000-09-26  
; PRIOR APPLICATION NUMBER: US/60/235,711  
; PRIOR FILING DATE: 2000-09-27  
; PRIOR APPLICATION NUMBER: US/60/235,720  
; PRIOR FILING DATE: 2000-09-27  
; PRIOR APPLICATION NUMBER: US/60/235,840  
; PRIOR FILING DATE: 2000-09-27  
; PRIOR APPLICATION NUMBER: US/60/235,863  
; NUMBER OF SEQ ID NOS: 2276  
; SOFTWARE: PatentIn version 3.0  
; SEQ ID NO 760  
; LENGTH: 1157  
; TYPE: DNA  
; ORGANISM: Homo sapiens  
US-09-954-456-760

Query Match 3.6%; Score 51.4; DB 10; Length 1157;  
Best Local Similarity 46.5%; Pred. No. 0.0073;  
Matches 166; Conservative 0; Mismatches 191; Indels 0; Gaps 0;

Qy 235 AAGATCCACCGCAGCAGCCGCGAGCTGATGTGCGAGTCTGCGAGGCGGAGCGC 294  
Db 488 ATGCTCGGCCAGAGCACCGGAGAGCTGCGGCTGCGCTTCCACTGCGAGAGCTG 547

QY	295	TGSGGCGGCTTCCTCCGCGCCCTACCGGAGGGGGGGCGCGCGCGCTGCAGAGAGC	354
Db	548	CGTAAGCGGGCTCTCCGGATCCGGATACCTGCAGAAAGCCCTGGGAGGTACAGGCC	607
QY	355	CTGGCGGCGCGGCTTCGCCAGCACTCGGTGCGCGCTGCAACTGAGAGCTGCGCGCGGCC	414
Db	608	GGGGGCCCCCGAGGGGCGCGAGCGGCTCTCAGGCCCATTCGCGAGGCGCTTGAGGCCCTTG	667
QY	415	GAGCGGCTGGAACGCTTTCCTGGCGGAGCGAGGAGGCTGTTTGATTCATCTCAGGCCAG	474
Db	668	GTCGAACAGGGGCGCGGTGCGGGCGGCCACTGTGGGCTTCCTGCGCGGCGACGCGCTACG	727
QY	475	CAGCCGACCGGCTCCGGATGAAGAACTGGCTGAGCTGAGAGATGCGCAATCTG	534
Db	728	GAGCGGGCCCGAGGCTCTGGGGGAGAGGCGCTGCGCGCGGAGTGAAGATGGGACGCGG	787
QY	535	AAGTCGGGCTGGGGGGCCCCGGGGGTGGCGACCGGGAGGTCTCTTGCGGCCCTTTCGAG	591
Db	788	ACCCCGACCGCTGACGAGAGTGAAGGAGGATGGGAGAGTTCGCGCGCAAGCTG	844

RESULT 6  
US-09-880-107-2244  
Sequence 2244, Application US/09880107  
Patent No. US20020142981A1  
GENERAL INFORMATION:  
APPLICANT: Horne, Darci T.  
APPLICANT: Vockley, Joseph G.  
APPLICANT: Scherf, Uwe  
APPLICANT: Gene Logic, Inc.  
TITLE OF INVENTION: Gene Expression Profiles in Liver Cancer  
FILE REFERENCE: 44921-5028-WO  
CURRENT APPLICATION NUMBER: US/09/880,107  
CURRENT FILING DATE: 2001-06-14  
PRIOR APPLICATION NUMBER: US 60/211,379  
PRIOR FILING DATE: 2000-06-14  
PRIOR APPLICATION NUMBER: US 60/237,054  
PRIOR FILING DATE: 2000-10-02  
NUMBER OF SEQ ID NOS: 3950  
SOFTWARE: PatentIn Ver. 2.1  
SEQ ID NO 2244  
LENGTH: 1157  
TYPE: DNA  
ORGANISM: Homo sapiens  
FEATURE:  
OTHER INFORMATION: Genbank Accession No. US20020142981A1 MI2529  
US-09-880-107-2244

Query Match	3.6%	Score 51.4;	DB 10;	Length 1157;
Best Local Similarity	46.5%	Pred. No. 0.0073;		
Matches 166;	Conservative 0;	Mismatches 191;	Indels 0;	Gaps 0

Qy	235	AAAGATCCACCGAGGACCCGAGCTATATCGTGGACGCTGCAATTCGCGGCGCGACGCC	294
Db	488	ATGCTCGGCAGAGACACCGAGAGACTGGGAGTGGCGCTTCGCTCCACTGCGCAGACTG <td>557</td>	557
Qy	295	TGTGGCGGCTTCTCTCCGCGCTTACCGCAGGGGCGCGCTGCGCGCGCGCTGCAGAGAGC <td>354</td>	354
Db	548	CGTAAAGCGGCTCTCTCCGAGATCCGATGACCTGCAGAAAGCGGCTGGCAGTGTACAGGCC <td>607</td>	607
Qy	355	CTGGCGGCGCGGCTCGGCCACACACTCGATGCGCGCTGCACATGTGAGGCTGCGCGCGCGGCC <td>414</td>	414
Db	608	GAGGCGCGCGAGGGGCGCGACCGCGGCGCTCAGCGCCATCGCGAGCGCGCTGGGGCGCCCTG <td>667</td>	667
Qy	415	GAGCGGCTGACGCTTTTGCTGTGGCGGACGAGAGCGCTGTTTGAATTGCATCTAGCCAG <td>474</td>	474
Db	668	GTAAGAACAGGGCGCGGTGCGGCGCGCACTGTGGGCTCCGCGCGGCGCAGCGCGTACAG <td>727</td>	727
Qy	475	CAGCCCAACCGGCTTCGGGATGAAGAAATTGGCTGAGCTGGAGAGATGGCGCTCGAAATCTG <td>534</td>	534
Db	728	GAGCGGCGCCAGGCGCTGGGGGCGAGCGGCTTGGCGCGCGAGATGAGAGAGATGGGCAGTGG <td>787</td>	787

Qy 535 AAGTGGCGCTCGGGGGCCCCGGGGTGGCCACGGGGAGGTGCTTTGGCCCCCTTGCAG 591

Db 788 ACCCGGACCGGCTTGACGAGGTGAAGAGACAGTGGCGAGGTGCGCGCAAGCTG 844

```

RESULT 7
US-09-967-013-5
: Sequence 5, Application US/09967013
: Patent No. US20020045840A1
: GENERAL INFORMATION:
: APPLICANT: Staton, Jr. Vincent P.
: TITLE OF INVENTION: METHOD FOR GENETIC AN
: TITLE OF INVENTION: AND USE THEREOF
: FILE REFERENCE: 11926-022001
: CURRENT APPLICATION NUMBER: US/09/967,013
: CURRENT FILING DATE: 2000-10-25
: PRIOR APPLICATION NUMBER: 60/206,613
: PRIOR FILING DATE: 2000-05-23
: NUMBER OF SEQ ID NOS: 91
: SOFTWARE: FASTSQ for Windows Version 4.0
: SEQ ID NO 5
: LENGTH: 41907
: TYPE: DNA
: ORGANISM: Homo sapiens
US-09-967-013-5

```

Query Match	3.6%	Score 51.2;	DB 10;	Length 41907;
Best Local Similarity	43.8%	Pred. No. 0.02;		
Matches 212;	Conservative 2;	Mismatches 270;	Indels 0;	Gaps 0;

QY	108	GTTGTCGTGGAAACAAGTGTGCTCTGTGGATGCTCAAGCGAATCCCGACGAAAGATGGC	167
Db	21162	GACCCGCGTGTGGGAAGAACCGGGACAGGCTGTCCAAAGAGCTTGAGGGCGCGGAGGC	21221
QY	168	AGGTACAGAGGCTCTGCAGGCTGCCTTTGGACAGAGCGGCGAGGCCGAGACCTGTGCTGCA	227
Db	21222	CCGGCTGTGGCGGGAATGGAAGACGTGTGGCCGCTCTGTGCATGATCCGCGGCAAGT	21281
QY	228	GATGCTGAAGATCCACCGCAGCGAACCCGACGTGATGTGACGTGCGATTCTGCGAGCG	287
Db	21282	GCAAGGCATGCTCGGCAAGACCAAGAGAACTGGGGTGGCCTCGCCTCCACCTGGG	21341
QY	288	GCAACCTGTGGCCGCTTCTTCGCGCTTACCGGAGAGGGCGCTGGCGCCGCGCTGCA	347
Db	21342	CAAGCTGTGAAGCGGCTCTCTCCGAGATCCGATGACCTGCAGAAAGCTCTGGCACTGTA	21401
QY	348	GAGAGCTGTGGCGCGCGCTGCGCCACACACTCGGTCCGCTGCACTGGAGCTGGCGCG	407
Db	21402	CCAGCGCGGGGCGCGGAGAGGGCGCGACCGGGCTTCAAGCCATCCGCGAGGCGCTGGG	21461
QY	408	CGGCGCGAGCGGCTTGACGCTTTTCTGTGGCGAAGAGAGCGCTGTGATGATCACT	467
Db	21462	GCCCCGTGTGAACAAGGGCGCGGTGCGGCGCCACTGTGGCTCCTTGGCCGGCCAGCC	21521
QY	468	AGCCCAAGCAGCCCGACCGGCTTCGGGATGAAGAACTGAGCTGTGAGGATGCGTGGG	527
Db	21522	GCTTACAGAGCGGGCCCAAGGCTGTGGGGGAGAGGGCTGTGGCGCGCGGATGGAAGATGGG	21581
QY	528	AAATCTGAAGTGGGCTCGGGGGCCCGGGGTGGCGAAGGGAGGTGCTTTCGCGCCCTT	587
Db	21582	CAGCGGACCCCGGACCGCTTGATGAGATGAAAGACAGAGTGAGGAGGTGGCGGCCAA	21641
QY	588	GCAG 591	
Db	21642	GCTG 21645	

RESULT 8  
US-09-870-759-129  
; Sequence 129, Application US/09870755  
; Patent No. US20020177551A1  
; GENERAL INFORMATION:  
; APPLICANT: TERMAN, David S

;; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR TREATMENT OF NEOPLASTIC DISEASE  
;; FILE REFERENCE: 870759  
;; CURRENT APPLICATION NUMBER: US/09/870,759  
;; CURRENT FILING DATE: 2002-01-14  
;; PRIOR APPLICATION NUMBER: US 60/208,128  
;; PRIOR FILING DATE: 2000-05-30  
;; NUMBER OF SEQ ID NOS: 166  
;; SOFTWARE: PatentIn version 3.1  
;; SEQ ID NO 129  
;; LENGTH: 1156  
;; TYPE: DNA  
;; ORGANISM: Homo sapiens  
;; FEATURE:  
;; NAME/KEY: CDS  
;; LOCATION: (61)..(1014)  
;; OTHER INFORMATION:  
US-09-870-759-129

Query Match 3.5%; Score 50.4; DB 9; Length 1156;  
Best Local Similarity 44.0%; Pred. No. 0.012;  
Matches 213; Conservative 0; Mismatches 271; Indels 0; Gaps 0;

QY 108 GTCTCTGCTGACAGAGTGTCTCTGTGATGCTTACCGCAGCCCGACGAGAAAGTGGC 167  
DB 360 GACCCCGGTGGCGAGAGACGCGGCGACGCTGTCCAGAGAGCTGCGAGCGCGCAGGC 419  
QY 168 AGGTACAGGGCTCTGACAGCTGCTTGGCAGAGAGCGGCGAGCCCGACGTCGCA 227  
DB 420 CCGGCTGGGGCGGACATGAGAGCTGTGGCGCGCTGTGCAATACCGCGCAGGT 479  
QY 228 GATGCTGAAGATCCACCGACGACCCCGACGCTGATGTCAGCTGCGATTTCGCGGCG 287  
DB 480 GCGGCGCATGCTGGGCGAGACCGAGAGCTGCGGCTGCGCTCCACCTCCAGCTGGC 539  
QY 288 GCGACCTGTGGCGCTTCTCTCGCGCTTACCGGAGGGGCGCTGCGCGCGCTGCA 347  
DB 540 CAGGCTGCGTAAGCGGCTCTCTCCGCGATGCGCATGACCTTCAGAGAGCGCTGCGAGTGA 599  
QY 348 GAGGAGCTGCGGCGCGCGCTGCGCGACCTGCGTCCGCTGCAATGAGAGCTGCGCGC 407  
DB 600 CAGGCGCGGCGCGCGCGCGCGCGCGCGCGCGCTCAGCGCCATCCGCGAGCGCTGGG 659  
QY 408 CCGGCGCGGCGCGCTGAGCGCTTGTCTGCGGAGACGAGAGCGCTTGTGATTCATCT 467  
DB 660 GCCCTGTGTGAACAGGGCGCGGCGCGCGCGCGCGCGCATGTGGGCTCCCTGGCGCGCAGCC 719  
QY 468 AGCCGACGACCGGCGCGCTGCGGCTGCGGATGAAGAACTGAGCTGAGGATGCGCTGGC 527  
DB 720 GCTACAGAGAGCGGCGCGCGCTGCGGCGCGAGCGCGCTGCGCGCGGATGAGAGATGGG 779  
QY 528 AAATCTGAAGTGGCGCTGCGGGGCGCGGGGATGAGAGCGGAGAGTCCCTTCGCGCCCTT 587  
DB 780 CAGCGGAGCCCGGACCGCTGCGAGGAGTGAAGAGCAGAGTGGCGAGGTTGCGCGCAA 839  
QY 588 GCAG 591  
DB 840 GCTG 843

RESULT 9  
US-09-827-854-7  
;; Sequence 7, Application US/09827854  
;; Patent No. US20020123093A1  
;; GENERAL INFORMATION:  
;; APPLICANT: Zannis, Vasile  
;; APPLICANT: Kyriacos, Kyriacos E.  
;; TITLE OF INVENTION: Compounds and methods for lowering  
;; TITLE OF INVENTION: cholesterol levels without inducing hypertriglyceridemia  
;; FILE REFERENCE: 07180/004003  
;; CURRENT APPLICATION NUMBER: US/09/827,854  
;; CURRENT FILING DATE: 2001-04-05  
;; PRIOR APPLICATION NUMBER: US 09/679,088  
;; PRIOR FILING DATE: 2000-10-04

;; PRIOR APPLICATION NUMBER: US 09/544,386  
;; PRIOR FILING DATE: 2000-04-06  
;; NUMBER OF SEQ ID NOS: 19  
;; SOFTWARE: FastSeq for Windows Version 4.0  
;; SEQ ID NO 7  
;; LENGTH: 1156  
;; TYPE: DNA  
;; ORGANISM: Homo sapiens  
US-09-827-854-7

Query Match 3.5%; Score 50.4; DB 10; Length 1156;  
Best Local Similarity 44.0%; Pred. No. 0.012;  
Matches 213; Conservative 0; Mismatches 271; Indels 0; Gaps 0;

QY 108 GTCTCTGCTGACAGAGTGTCTCTGTGATGCTTACCGCAGCCCGACGAGAAAGTGGC 167  
DB 360 GACCCCGGTGGCGAGAGACGCGGCGACGCTGTCCAGAGAGCTGCGAGCGCGCAGGC 419  
QY 168 AGGTACAGGGCTCTGACAGCTGCTTGGCAGAGAGCGGCGAGCCCGACGTCGCA 227  
DB 420 CCGGCTGGGGCGGACATGAGAGCTGTGGCGCGCTGTGCAATACCGCGCAGGT 479  
QY 228 GATGCTGAAGATCCACCGACGACCCCGACGCTGATGTCAGCTGCGATTTCGCGGCG 287  
DB 480 GCGGCGCATGCTGGGCGAGACCGAGAGCTGCGGCTGCGCTCCACCTCCAGCTGGC 539  
QY 288 GCGACCTGTGGCGCTTCTCTCGCGCTTACCGGAGGGGCGCTGCGCGCGCTGCA 347  
DB 540 CAGGCTGCGTAAGCGGCTCTCTCCGCGATGCGCATGACCTTCAGAGAGCGCTGCGAGTGA 599  
QY 348 GAGGAGCTGCGGCGCGCGCTGCGCGACCTGCGTCCGCTGCAATGAGAGCTGCGCGC 407  
DB 600 CAGGCGCGGCGCGCGCGCGCGCGCGCGCGCGCTCAGCGCCATCCGCGAGCGCTGGG 659  
QY 408 CCGGCGCGGCGCGCTGAGCGCTTGTCTGCGGAGACGAGAGCGCTTGTGATTCATCT 467  
DB 660 GCCCTGTGTGAACAGGGCGCGGCGCGCGCGCGCGCGCATGTGGGCTCCCTGGCGCGCAGCC 719  
QY 468 AGCCGACGACCGGCGCGCTGCGGATGAAGAACTGAGCTGAGAGATGCGCTGGC 527  
DB 720 GCTACAGAGAGCGGCGCGCGCTGCGGCGCGAGCGCTGCGCGCGGATGAGAGATGGG 779  
QY 528 AAATCTGAAGTGGCGCTGCGGGGCGCGGGGATGAGAGCGGAGAGTCCCTTCGCGCCCTT 587  
DB 780 CAGCGGAGCCCGGACCGCTGCGAGGAGTGAAGAGCAGAGTGGCGAGGTTGCGCGCAA 839  
QY 588 GCAG 591  
DB 840 GCTG 843

RESULT 10  
US-09-827-854-8  
;; Sequence 8, Application US/09827854  
;; Patent No. US20020123093A1  
;; GENERAL INFORMATION:  
;; APPLICANT: Zannis, Vasile  
;; APPLICANT: Kyriacos, Kyriacos E.  
;; TITLE OF INVENTION: Compounds and methods for lowering  
;; TITLE OF INVENTION: cholesterol levels without inducing hypertriglyceridemia  
;; FILE REFERENCE: 07180/004003  
;; CURRENT APPLICATION NUMBER: US/09/827,854  
;; CURRENT FILING DATE: 2001-04-05  
;; PRIOR APPLICATION NUMBER: US 09/679,088  
;; PRIOR FILING DATE: 2000-10-04  
;; PRIOR APPLICATION NUMBER: US 09/544,386  
;; PRIOR FILING DATE: 2000-04-06  
;; NUMBER OF SEQ ID NOS: 19  
;; SOFTWARE: FastSeq for Windows Version 4.0  
;; SEQ ID NO 8  
;; LENGTH: 1156  
;; TYPE: DNA  
;; ORGANISM: Homo sapiens

US-09-827-854-8

Query Match 3.5%; Score 50.4; DB 10; Length 1156;  
Best Local Similarity 44.0%; Pred. No. 0.012;  
Matches 213; Conservative 0; Mismatches 271; Indels 0; Gaps 0;

QY 108 GTCCTGCTGACACAGAGTGTCTGTGAGATGCTTACCGGCAACCCCAAGAGTGGC 167  
DB 360 GACCCCGGTGGGAGAGAGACGCGGGCAAGCTGTCAAGAGCTGACAGGCGCGCAGGC 419  
QY 168 AGTGTACAGGGGCTCTGACAGGCTGCTTGCCAGAGCGGGGAGAACCCGAGAGTGTGCA 227  
DB 420 CCGGCTGGGCGCGGACATGAGAGAGCTGTGCGCGGCTGTGTGACATGACCGCGGAGGT 479  
QY 228 GATGCTGAAGATTCACCCGACGACCCGACGCTGATGTGACAGTGTGCAATTTCTGGGAGC 287  
DB 480 GCAGGCCATGCTCTGGCCAGAGACCGAGAGAGTGTGGGGTGGCCCTGCTCCACCTGGC 539  
QY 288 GCAAGCTGTGGCGGCTTCTTCGCGCTTACCGGAGGGGGCGCTGGCGCGCGCTGCA 347  
DB 540 CAAGCTGCTGAAGCGGCTTCTTCGCGATGCGATGACCTGTGAGAAAGCGCTGGCAGTGA 599  
QY 348 GAGAGGCTGTGGGCGCGGCTGCGCCAGACATGCGGCGGCTGCAACTGTGAGAGTGTGCGC 407  
DB 600 CCAAGCGGGGCGCGGAGAGGCGCGGAGCGGCTTCAAGCGCATTCGCGAGCGCGCTGG 659  
QY 408 CGGCGCCGAGCGGCTGACGCTTTGCTGTGGCGGACGAGAGCGGCTGTGATGATCTCT 467  
DB 660 GCCCTGTGTGAACAGAGGCGCGGCTGTGGGCGCGCACTGTGGGCTCCCTGGCGGCAAGC 719  
QY 468 AGCCAGCAGACCCGACCGGCTTCGCGATGAGAACTGTGAGTGTGAGAGTGTGCG 527  
DB 720 GCTACAGAGCGGCGCCAGAGGCTGTGGGCGGAGCGGCTGCGCGGATGAGAGATGG 779  
QY 528 AAATGTGAAGTGGGCTTCTTCGCGGCTGCGGCGGCTGCGGAGTGTGCTTGGCCCTT 587  
DB 780 CAGCGGAGCCGCGGACCGCTTGAAGAGTGAAGAGCAGGTGGGAGAGTGTGCGGCCA 839  
QY 588 GCAG 591  
DB 840 GCTG 843

RESULT 11  
US-09-827-854-10  
; Sequence 10, Application US/09827854  
; Patent No. US20020123093A1  
; GENERAL INFORMATION:  
; APPLICANT: Zannis, Vassilis  
; APPLICANT: Kypreos, Kyriakos E.  
; TITLE OF INVENTION: Compounds and methods for lowering  
; FILE REFERENCE: 07180/004003  
; CURRENT APPLICATION NUMBER: US/09/827,854  
; PRIOR FILING DATE: 2001-04-05  
; PRIOR APPLICATION NUMBER: US 09/679,088  
; PRIOR FILING DATE: 2000-10-04  
; PRIOR APPLICATION NUMBER: US 09/544,386  
; NUMBER OF SEQ ID NOS: 19  
; SOFTWARE: FastSeq for Windows Version 4.0  
; SEQ ID NO 10  
; LENGTH: 1156  
; TYPE: DNA  
; ORGANISM: Homo sapiens  
US-09-827-854-10

Query Match 3.5%; Score 50.4; DB 10; Length 1156;  
Best Local Similarity 44.0%; Pred. No. 0.012;  
Matches 213; Conservative 0; Mismatches 271; Indels 0; Gaps 0;

QY 108 GTCCTGCTGACACAGTGTCTGTGAGATGCTTACCGGCAACCCCAAGAGTGGC 167  
DB 360 GACCCCGGTGGGAGAGAGACGCGGGCAAGCTGTCAAGAGCTGACAGGCGCGCAGGC 419  
QY 168 AGTGTACAGGGGCTCTGACAGGCTGCTTGCCAGAGCGGGGAGAACCCGAGAGTGTGCA 227  
DB 420 CCGGCTGGGCGCGGACATGAGAGAGCTGTGCGCGGCTGTGTGACATGACCGCGGAGGT 479  
QY 228 GATGCTGAAGATTCACCCGACGACCCGACGCTGATGTGACAGTGTGCAATTTCTGGGAGC 287  
DB 480 GCAGGCCATGCTCTGGCCAGAGACCGAGAGAGTGTGGGGTGGCCCTGCTCCACCTGGC 539  
QY 288 GCAAGCTGTGGCGGCTTCTTCGCGCTTACCGGAGGGGGCGCTGGCGCGCGCTGCA 347  
DB 540 CAAGCTGCTGAAGCGGCTTCTTCGCGATGCGATGACCTGTGAGAAAGCGCTGGCAGTGA 599  
QY 348 GAGAGGCTGTGGGCGCGGCTGCGCCAGACATGCGGCGGCTGCAACTGTGAGAGTGTGCGC 407  
DB 600 CCAAGCGGGGCGCGGAGAGGCGCGGAGCGGCTTCAAGCGCATTCGCGAGCGCGCTGG 659  
QY 408 CGGCGCCGAGCGGCTGACGCTTTGCTGTGGCGGACGAGAGGCGCTGTGATGATCTCT 467  
DB 660 GCCCTGTGTGAACAGAGGCGCGGCTGTGGGCGCGCACTGTGGGCTCCCTGGCGGCAAGC 719  
QY 468 AGCCAGCAGCCGACCGGCTTCGCGATGAGAACTGTGAGTGTGAGAGTGTGCG 527  
DB 720 GCTACAGAGCGGCGCCAGAGGCTGTGGGCGGAGCGGCTGCGCGGATGAGAGATGG 779  
QY 528 AAATGTGAAGTGGGCTTCTTCGCGGCTGCGGCGGCTGCGGAGTGTGCTTGGCCCTT 587  
DB 780 CAGCGGAGCCGCGGACCGCTTGAAGAGTGAAGAGCAGGTGGGAGAGTGTGCGGCCA 839  
QY 588 GCAG 591  
DB 840 GCTG 843

DB 360 GACCCCGGTGGGAGAGAGACGCGGGCAAGCTGTCAAGAGCTGACAGGCGCGCAGGC 419  
QY 168 AGTGTACAGGGGCTCTGACAGGCTGCTTGCCAGAGCGGGGAGAACCCGAGAGTGTGCA 227  
DB 420 CCGGCTGGGCGCGGACATGAGAGAGCTGTGCGCGGCTGTGTGACATGACCGCGGAGGT 479  
QY 228 GATGCTGAAGATTCACCCGACGACCCGACGCTGATGTGACAGTGTGCAATTTCTGGGAGC 287  
DB 480 GCAGGCCATGCTCTGGCCAGAGACCGAGAGAGTGTGGGGTGGCCCTGCTCCACCTGGC 539  
QY 288 GCAAGCTGTGGCGGCTTCTTCGCGGCTTACCGGAGGGGGCGCTGGCGCGCGCTGCA 347  
DB 540 CAAGCTGCTGAAGCGGCTTCTTCGCGATGCGATGACCTGTGAGAAAGTGTGCGGAGTGA 599  
QY 348 GAGAGGCTGTGGGCGCGGCTGCGCCAGACATGCGGCGGCTGCAACTGTGAGAGTGTGCGC 407  
DB 600 CCAAGCGGGGCGCGGAGAGGCGCGGAGCGGCTTCAAGCGCATTCGCGAGCGCGCTGG 659  
QY 408 CGGCGCCGAGCGGCTGACGCTTTGCTGTGGCGGACGAGAGGCGCTGTGATGATCTCT 467  
DB 660 GCCCTGTGTGAACAGAGGCGCGGCTGTGGGCGCGCACTGTGGGCTCCCTGGCGGCAAGC 719  
QY 468 AGCCAGCAGCCGACCGGCTTCGCGATGAGAACTGTGAGTGTGAGAGTGTGCG 527  
DB 720 GCTACAGAGCGGCGCCAGAGGCTGTGGGCGGAGCGGCTGCGCGGATGAGAGATGG 779  
QY 528 AAATGTGAAGTGGGCTTCTTCGCGGCTGCGGCGGCTGCGGAGTGTGCTTGGCCCTT 587  
DB 780 CAGCGGAGCCGCGGACCGCTTGAAGAGTGAAGAGCAGGTGGGAGAGTGTGCGGCCA 839  
QY 588 GCAG 591  
DB 840 GCTG 843

RESULT 12  
US-09-827-854-12  
; Sequence 12, Application US/09827854  
; Patent No. US20020123093A1  
; GENERAL INFORMATION:  
; APPLICANT: Zannis, Vassilis  
; APPLICANT: Kypreos, Kyriakos E.  
; TITLE OF INVENTION: Compounds and methods for lowering  
; FILE REFERENCE: 07180/004003  
; CURRENT APPLICATION NUMBER: US/09/827,854  
; PRIOR FILING DATE: 2001-04-05  
; PRIOR APPLICATION NUMBER: US 09/679,088  
; PRIOR FILING DATE: 2000-10-04  
; PRIOR APPLICATION NUMBER: US 09/544,386  
; NUMBER OF SEQ ID NOS: 19  
; SOFTWARE: FastSeq for Windows Version 4.0  
; SEQ ID NO 12  
; LENGTH: 1156  
; TYPE: DNA  
; ORGANISM: Homo sapiens  
US-09-827-854-12

Query Match 3.5%; Score 50.4; DB 10; Length 1156;  
Best Local Similarity 44.0%; Pred. No. 0.012;  
Matches 213; Conservative 0; Mismatches 271; Indels 0; Gaps 0;

QY 108 GTCCTGCTGACACAGTGTCTGTGAGATGCTTACCGGCAACCCCAAGAGTGGC 167  
DB 360 GACCCCGGTGGGAGAGAGACGCGGGCAAGCTGTCAAGAGCTGACAGGCGCGCAGGC 419  
QY 168 AGTGTACAGGGGCTCTGACAGGCTGCTTGCCAGAGCGGGGAGAACCCGAGAGTGTGCA 227  
DB 420 CCGGCTGGGCGCGGACATGAGAGAGCTGTGCGCGGCTGTGTGAGATGACCGCGGAGGT 479  
QY 228 GATGCTGAAGATTCACCCGACGACCCGACGCTGATGTGACAGTGTGCAATTTCTGGGAGC 287  
DB 480 GCAGGCCATGCTCTGGCCAGAGACCGAGAGAGTGTGGGGTGGCCCTGCTCCACCTGGC 539  
QY 288 GCAAGCTGTGGCGGCTTCTTCGCGCTTACCGGAGGGGGCGCTGGCGCGCGCTGCA 347  
DB 540 CAAGCTGCTGAAGCGGCTTCTTCGCGATGCGATGACCTGTGAGAAAGTGTGCGGAGTGA 599  
QY 348 GAGAGGCTGTGGGCGCGGCTGCGCCAGACATGCGGCGGCTGCAACTGTGAGAGTGTGCGC 407  
DB 600 CCAAGCGGGGCGCGGAGAGGCGCGGAGCGGCTTCAAGCGCATTCGCGAGCGCGCTGG 659  
QY 408 CGGCGCCGAGCGGCTGACGCTTTGCTGTGGCGGACGAGAGGCGCTGTGATGATCTCT 467  
DB 660 GCCCTGTGTGAACAGAGGCGCGGCTGTGGGCGCGCACTGTGGGCTCCCTGGCGGCAAGC 719  
QY 468 AGCCAGCAGCCGACCGGCTTCGCGATGAGAACTGTGAGTGTGAGAGTGTGCG 527  
DB 720 GCTACAGAGCGGCGCCAGAGGCTGTGGGCGGAGCGGCTGCGCGGATGAGAGATGG 779  
QY 528 AAATGTGAAGTGGGCTTCTTCGCGGCTGCGGCGGCTGCGGAGTGTGCTTGGCCCTT 587  
DB 780 CAGCGGAGCCGCGGACCGCTTGAAGAGTGAAGAGCAGGTGGGAGAGTGTGCGGCCA 839  
QY 588 GCAG 591  
DB 840 GCTG 843

Db	480	GCAGGCCATGCTCGGCGCCAGAGACACCGAGAGACTGCGGGGTGGCGCTCGCGCTCCACCTGCG	53.9
Qy	288	GCAGCCCTGTGTGGCGGCTTCTCTCGCGGCTTACCGGAGGGGGCGCTGGCGCGCCGCTGCA	34.7
Db	540	CAGGTTGTGTGTCAGCGCGCTCTCTCCGCGATGCGGATGACTGTGCAGAAAGCGCGCTGGCAGGTGA	53.9
Qy	348	GAGGAGCGCTGGCGGGCGGCGCTCGGCCACGACTCGGTGCTGTGAACGTGAGAGCTGCGGCGC	40.7
Db	600	CCAGGCCCGGGGCGCGCGAGGGGGCGGACCGCGGCTTCAAGCGCCATTCGCGAGCGCCTGGG	65.9
Qy	408	CGGCGCCGAGCGGCTGTGACGCGTTTGTCTGGCGGAGCGAGGAGCGCTGTTTGAAGTTGCATCT	46.7
Db	660	GCCCCGTGTGGAACAGGGGCCGCGGTGTGCGGCGCGCACACTGTGGCTTCCCTGGCCGCGCCAGCC	71.9
Qy	468	AGCCACGACGCCCGACCGGCTTCCGGGATGAAAGACTGCGTGAAGGTGAGATGCGCTGCG	52.7
Db	720	GCTACAGAGAGCGGGCGCCAGGCGCTTGGGGCGAGCGGCTGCGCGCGGATGAGGAGATGGG	77.9
Qy	528	AAATCTGAATGTGCGGCTTCGGGGGGCCCCGGGGGTGGCGGACCGGGGAGAGTGCCTTCGCGCCCTT	58.7
Db	780	CAGCGCGACCGCGCGACCGCGCTTGGACGAGGTGAAGAGAGACAGGTGTGGCGGAGGTGCGCGCAA	83.9
Qy	588	GCAG	59.1
Db	840	GCTG	84.3

```

RESULT 13
US-10-044-090-454
; Sequence 454, Application US/10044090
; Patent No. US20020137081A1
; GENERAL INFORMATION:
; APPLICANT: Olga Bandman
; TITLE OF INVENTION: GENES DIFFERENTIALLY EXPRESSED IN VASCULAR TISSUE ACTIVATION
; FILE REFERENCE: PN-0028 US
; CURRENT APPLICATION NUMBER: US/10/044,090
; CURRENT FILING DATE: 2002-01-09
; NUMBER OF SEQ ID NOS: 850
; SOFTWARE: PERL Program
; SEQ ID NO 454
; LENGTH: 1291
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: misc feature
; OTHER INFORMATION: Incyte ID No. US20020137081A1 2514988CB1
; NAME/KEY: unsure
; LOCATION: 46
; OTHER INFORMATION: a, t, c, g, or other
US-10-044-090-454

Query Match          3.5%; Score 50.4; DB 12; Length 1291;
Best Local Similarity 44.0%; Pred. No. 0.013;
Matches 213; Conservative 0; Mismatches 271; Indels 0; Gaps 0;

Qy      108  GTCTCTGCTGACAAGTGTGTCCTGTGCGATGCTTACGCGCACCCCGACGACAAGTGGC 167
Db      485  GACCCCGGTGGCGGAGGAGACGCGGGCACCGCTGTCCAAAGAGCTGACGGCGGCGCAGGG 544

Qy      168  AGTGTACAGGGCTCTGCAAGGCTGCTTGGCAGAGAGCGCGGAGACCCGACGCTGTGCA 227
Db      545  CCGGTGGGCCCGACATGAGAGACGTGTGGGCGCGCTGTGATGACCGCGCGAGGT 604

Qy      228  GATGCTGAAGATTCACCGCAGCGACCCGCAAGCTGATGTCGACGTGATTTGCGGGGG 287
Db      605  GCAAGCCATGCTCGGCCACAGACACCGAGAGACTGTGGGTGGCTTCGCTTCCACCTGGG 664

Qy      288  GCAGCCCTGTGACCGCTTCTCCGCGGCTTACCGCGAGAGGGAGCGCTGGCGCGCGCTGCA 347
Db      665  CAAAGTGTGAAGCGGCTCTCCGCGATGCCGATGACCTGCAAGAGGCGCTGGCAGATGA 724

Qy      348  GAGAGAGCTGGCGGCGCGCTGCGCCACGACTCGGTCCTGCACACTGGAAGCTGGCGCG 407

```

Db	725	CGAAGCCGAGGCGCCGACAGAGGCGCCGAGAGCGGCTCAGAGGCCATCCGCGAAGCGGCTGGG	784
Qy	408	CGGCGCCCGAGCGCGCTGGACGCTTTGCTGCGCGAGCAGAGAGCGCTGTTTGAATTGCATCT	467
Db	785	GCCCTTGTTGGACAGAGGCGCGCGCGCGCCGACACTGTGGGCTCCCTGCGCGGCCAGCC	844
Qy	468	AGCCACGACAGCCCGACCGGCTCCGGGATGGAAGAACTGGAGCTGGAGGATCGCTGCG	527
Db	845	GCTACAGGAGCGGGGCCCAAGGCTCTGGGGCGAAGCGGCTGCGGCGCGGAGATGAAGAGATGGG	904
Qy	528	AAATCTGAAGTGGCGGCTCTGGGGGCGCCCGGGGTGGCGACGAGGAGGTGCTTTCGCGCCCTTT	587
Db	905	CAGCGGAGCCCGCAGCCGCTTGACGAGATGGAAGAGACAGTGGCGGAGGCTGGCGGCCAA	964
Qy	588	GCAG	591
Db	965	GCTG	968

```

RESULT 14
US-09-772-304A-1
; Sequence 1, Application US/09772304A
; Patent No. US20020146785A1
; GENERAL INFORMATION:
; APPLICANT: Mahishi, L.H.
; APPLICANT: Tripathi, G.
; APPLICANT: Ramchander, T.V.N.
; APPLICANT: Rawal, S.K.
; TITLE OF INVENTION: NOVEL ESCHERICHIA COLI HAVING ACCESSION
; TITLE OF INVENTION: NO. PTA 1579 AND ITS USE TO PRODUCE POLYHYDROXYBUTRATE
; FILE REFERENCE: A33943 066123.0103
; CURRENT APPLICATION NUMBER: US/09/772.304A
; CURRENT FILING DATE: 2001-01-29
; NUMBER OF SEQ ID NOS: 1
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 1
; LENGTH: 4826
; TYPE: DNA
; ORGANISM: Streptomyces aureofaciens
US-09-772-304A-1

Query Match          3.5%; Score 50.4; DB 10; Length 4826;
Best Local Similarity 42.9%; Pred. No. 0.018;
Matches 368; Conservative 0; Mismatches 481; Indels 9; Gaps 2

QY 103 GTGAGATCTTCGCTGGACAGAGTGTCTCTGTGCGATGCTTAACGCGACACCCCGACAGAAAG 162
DB 1377 GTGGCCGGGCGCGTGAAGCGCGGCGCTGTCTGTCTGCGCTCCGAGGGGGCCCTGTAAACGG 1436

QY 163 GTGGCAGGTACAGGGGCTCTGCAAGCTGCTTTGGCAGAGAGCGCGGGAGCCCGAAGCTG 222
DB 1437 GTGTCGCGCGCGCTGGCGCGCGCTGTCTGTCGCGGGCGCGCTTCGCGGCGCGGCCCGGAGTG 1496

QY 223 CTGCAGATGCTGAAGATCCACCGCAGGACCCCGCAGCTGATGTGTCACGTGGAATTTC 282
DB 1497 CCGGCTCGGCGTTCGCGCGGTGCGCGCGGTGTCTGGGACACCGTATATACAGGGGTTGGCG 1556

QY 283 GGGCGGCGAGCCCTTGCGGCGCGCTTCTCTCGCGCGCTAACCGCAGAGGGGCGCTGCGCGCCGCG 342
DB 1557 GCGCGGTCCTTGGGGGCTGCCCCGACAGAGCGCGCCCGCTGATGTCTGTGTGCGCGCG 1616

QY 343 CTGCAGAGAGACTTGAGCGGCGCGCGCTTCGCCACAGCACTCGATGCGCGCTCAACTGAGCTG 402
DB 1617 CTGACGGGTGCTGTGCGGGGACAGACTGCGGGCGGAACCCGGTGGCGCTTCGCCCGTGAAGCTG 1676

QY 403 CGCGCGCGCGCGCGAGCGGCTGGAAGCTTTTGTCTGGCGGAGACAGAGAGCGCTGTTTGAAGTTGC 462
DB 1677 ACGGAGGCGTGGGCGCGCTGTGTGCGCCGATGGAGGCTCGGAGGTGTCCGAAGCGCGCGCTTTCG 1736

QY 463 ATCTTAGCCCAAGCAGACCCGACCGGCTCGGGATGAAGAACTGGAGCTGAGCTGAGAGGATGCG 522
DB 1737 GCCCTGTGCTTGGCGCTGTGTGCGGGGTGCTGGCGGCGGGGAGCGCGCGGTGTGCGCGCGGAGCGG 1796

```

Qy 523 CTGCGAAATCTGAACTGCGGCTCGGGGGCCGGGGTGCGACGGGAGAGTCGCTTCGGCC 582  
Db 1797 CGGCGCGAGGGGGGCTACTGCGCGCGACGCTG-----TCGTGCTGGCCACGTGGGG 1850  
Qy 583 CCCTTGACGCCCCCGGTGCTCTCTGTGCGAGTGAAACCGCCGCCGCCGCTT 642  
Db 1851 CGGCTGGGGGCTCGAGAGTGTGCTTCCGGAGGCGGTACACGCTCCGATGACGGTGGCC 1910  
Qy 643 GCCGACATTTTCTGTTCAGGGGTGAGCCTGATGTAATCGGCGGCTGAGCCTGAAGAC 702  
Db 1911 GCGCTGCTGTGTCGCGCGCGCGCGCGCGCGCGGACCCGAGAGCTCTGTCTGAGACGGCG 1970  
Qy 703 CAACAGACGTTTCGCGCGCTGTGTGGATCTCAATGACGCGAGAGTGGGGCTCACTGACG 762  
Db 1971 TACGGGGCGG---GGGCTGGCGCGCGACGCTGCTGCCAGCTGGCGGTGGCTTGACCGAC 2027  
Qy 763 CGAGGCTGCGGGGCGCTGCGGAGCCCGGCGCTGGAATCGTGGCTTACAGATGACAGCGC 822  
Db 2028 CCGGACTGCTCAGGCGCGCTGTGCTGGGGACGGCGCGCTGTGATCACCCTGCTCGGC 2087  
Qy 823 GAGGACTGTACGAGACGGGCTTCAGGCTGTCGGCGCTTGCGAGGCGGAGGCGCGC 882  
Db 2088 GCGCGCCACCGGCTTCAGGCGCTGCTGCTGCTGCGGCGGACGCTGTGCACTGGCTGGC 2147  
Qy 883 CGCGCCACGCTGACGCGCTGTGTGGAGGCACTCGAGAGAACGAGCTCACGCTGGCA 942  
Db 2148 CTGACACGAGCTGCGCCCGTACGCTGTGACAGATCGGGGTGGCCTCCCGCGTGGCTCCG 2207  
Qy 943 GAGACTGTGCTGGCGCTG 960  
Db 2208 CCGCGCTGCGCGGCTG 2225

## RESULT 15

US-09-976-059-1  
; Sequence 1, Application US/09976059  
; Patent No. US20020164747A1  
; GENERAL INFORMATION:  
; APPLICANT: Farnet, Chris  
; APPLICANT: Zazopoulos, Emmanuel  
; APPLICANT: Stafia, Alfredo  
; TITLE OF INVENTION: Genes and Proteins for Biosynthesis of Ramoplanin  
; FILE REFERENCE: 3019-PC1  
; CURRENT APPLICATION NUMBER: US/09/976,059  
; CURRENT FILING DATE: 2001-10-15  
; NUMBER OF SEQ ID NOS: 34  
; SOFTWARE: PatentIn version 3.0  
; SEQ ID NO 1  
; LENGTH: 88421  
; TYPE: DNA  
; ORGANISM: Actinoplanes sp.  
; FEATURE:  
; NAME/KEY: misc\_feature  
; LOCATION: (2077)..(3078)  
; OTHER INFORMATION: ORF 1; positive strandedness  
; NAME/KEY: misc\_feature  
; LOCATION: (3118)..(4032)  
; OTHER INFORMATION: ORF 2; positive strandedness  
; NAME/KEY: misc\_feature  
; LOCATION: (4038)..(5048)  
; OTHER INFORMATION: ORF 3; positive strandedness  
; NAME/KEY: misc\_feature  
; LOCATION: (4038)..(5048)  
; OTHER INFORMATION: ORF 3; positive strandedness  
; NAME/KEY: misc\_feature  
; LOCATION: (6665)..(5814)  
; OTHER INFORMATION: ORF 4; negative strandedness  
; NAME/KEY: misc\_feature  
; LOCATION: (7703)..(6693)  
; OTHER INFORMATION: ORF 5; negative strandedness  
; NAME/KEY: misc\_feature  
; LOCATION: (9464)..(8130)  
; OTHER INFORMATION: ORF 6; negative strandedness

; NAME/KEY: misc\_feature  
; LOCATION: (9691)..(10761)  
; OTHER INFORMATION: ORF 7; positive strandedness  
; NAME/KEY: misc\_feature  
; LOCATION: (12751)..(10829)  
; OTHER INFORMATION: ORF 8; negative strandedness  
; NAME/KEY: misc\_feature  
; LOCATION: (13617)..(12802)  
; OTHER INFORMATION: ORF 9; negative strandedness  
; NAME/KEY: misc\_feature  
; LOCATION: (15203)..(13614)  
; OTHER INFORMATION: ORF 10; negative strandedness  
; NAME/KEY: misc\_feature  
; LOCATION: (15591)..(15863)  
; OTHER INFORMATION: ORF 11; positive strandedness  
; NAME/KEY: misc\_feature  
; LOCATION: (15860)..(19035)  
; OTHER INFORMATION: ORF 12; positive strandedness  
; NAME/KEY: misc\_feature  
; LOCATION: (19032)..(39713)  
; OTHER INFORMATION: ORF 13; positive strandedness  
; NAME/KEY: misc\_feature  
; LOCATION: (39713)..(65800)  
; OTHER INFORMATION: ORF 14; positive strandedness  
; NAME/KEY: misc\_feature  
; LOCATION: (65826)..(66530)  
; OTHER INFORMATION: ORF 15; positive strandedness  
; NAME/KEY: misc\_feature  
; LOCATION: (66546)..(67370)  
; OTHER INFORMATION: ORF 16; positive strandedness  
; NAME/KEY: misc\_feature  
; LOCATION: (67384)..(70059)  
; OTHER INFORMATION: ORF 17; positive strandedness  
; NAME/KEY: misc\_feature  
; LOCATION: (70059)..(70662)  
; OTHER INFORMATION: ORF 18; positive strandedness  
; NAME/KEY: misc\_feature  
; LOCATION: (70659)..(71906)  
; OTHER INFORMATION: ORF 19; positive strandedness  
; NAME/KEY: misc\_feature  
; LOCATION: (73439)..(71964)  
; OTHER INFORMATION: ORF 20; negative strandedness  
; NAME/KEY: misc\_feature  
; LOCATION: (74216)..(73563)  
; OTHER INFORMATION: ORF 21; negative strandedness  
; NAME/KEY: misc\_feature  
; LOCATION: (75424)..(74213)  
; OTHER INFORMATION: ORF 22; negative strandedness  
; NAME/KEY: misc\_feature  
; LOCATION: (75535)..(76464)  
; OTHER INFORMATION: ORF 23; positive strandedness  
; NAME/KEY: misc\_feature  
; LOCATION: (78110)..(76449)  
; OTHER INFORMATION: ORF 24; negative strandedness  
; NAME/KEY: misc\_feature  
; LOCATION: (79864)..(78107)  
; OTHER INFORMATION: ORF 25; negative strandedness  
; NAME/KEY: misc\_feature  
; LOCATION: (81624)..(79861)  
; OTHER INFORMATION: ORF 26; negative strandedness  
; NAME/KEY: misc\_feature  
; LOCATION: (81909)..(81682)  
; OTHER INFORMATION: ORF 27; negative strandedness  
; NAME/KEY: misc\_feature  
; LOCATION: (82346)..(82062)  
; OTHER INFORMATION: ORF 28; negative strandedness  
; NAME/KEY: misc\_feature  
; LOCATION: (82587)..(84446)  
; OTHER INFORMATION: ORF 29; positive strandedness  
; NAME/KEY: misc\_feature  
; LOCATION: (84481)..(85548)  
; OTHER INFORMATION: ORF 30; positive strandedness  
; NAME/KEY: misc\_feature

; LOCATION: (85556)..(86845)  
; OTHER INFORMATION: ORF 31; positive strandedness  
; NAME/KEY: misc\_feature  
; LOCATION: (87372)..(86803)  
; OTHER INFORMATION: ORF 32; positive strandedness  
; NAME/KEY: misc\_feature  
; LOCATION: (87494)..(88420)  
; OTHER INFORMATION: ORF 33; positive strandedness; N-terminus only  
US-09-976-059-1

Query Match 3.5%; Score 50.2; DB 9; Length 88421;  
Best Local Similarity 46.4%; Pred. No. 0.041;  
Matches 163; Conservative 0; Mismatches 188; Indels 0; Gaps 0;

QY 97 CTGTTTGAGAGTCTCTGACAGAGTGTCTCTGCGATGCTTACGGACCCCGAG 156  
DB 65955 CTGCTCGACGACCTGACGCGGACGTCGTCTGCGGCGCGGACCTGCGCTTACCGTC 66014  
QY 157 CAGAGGTGGCAGTGTACAGGCTCTGACAGCTGCTTGCGAGAGCGGCGGAGCCCG 216  
DB 66015 GCCGACGAGGTGGCGGGGGGGCGGGGTGGCGCTTTGGGCAACGCTCGGGCGCGTC 66074  
QY 217 GACGTCTGACAGTGTGAAGATCACCGACGACCCGACGCTGATCTGCACTGCGA 276  
DB 66075 CTGCGGTACGAGATGCGCACGCGCTGTGTGCGACGCGAGGTGGCGTGAACGGCTC 66134  
QY 277 TTCTGCGGGGCGACCGCTGTGGCGCTTCTCGGCGCTTACCGGAGGGGGCGCTGCGC 336  
DB 66135 TTCTGACGGGCTGTGCGGATCTTGACCCCTTGACACACCGGCGAGGGCGCTTGAC 66194  
QY 337 GCCGCGCTGACAGAGACCTGCGGCGCGCTGCGCCAGCACTGCTGCGCGCTGCAACTG 396  
DB 66195 GACGAGAGATTCTGCTGCGGGTGGCGAGTTGCGCGTTACGACCAAGAGCGCTCGCC 66254  
QY 397 GACGTGCGCGCGCGCGCGCGCTGAGCGCTTGTCTGCGGAGCGAGAG 447  
DB 66255 GATCCGACATGCGCGAGCTGATCTGCGCGCTGCGCGCGACGTCGAG 66305

Search completed: February 3, 2003, 16:17:35  
Job time : 194 secs



**THIS PAGE BLANK (USPTO)**